FINNISH ROAD ADMINISTRATION

S12 Solutions to improve main roads

R. Tapio Luttinen, Riku Nevala

Capacity and Level of Service of Finnish Signalized Intersections

Finnra Reports 25/2002

S 12 Solutions to improve main roads

R. Tapio Luttinen, Riku Nevala

Capacity and Level of Service of Finnish Signalized Intersections

Finnra Reports

25/2002

Finnish Road Administration

Helsinki 2002

ISBN 951-726-903-X
ISSN 1457-9871
TIEH 3200757E

Oy Edita Ab

The publication is sold by:
Finnra, Publication Sales
Telefax: Int. +358 204 22 2652
Email: julkaisumyynti@Tiehallinto.fi
www.Tiehallinto.fi/julk2.htm

Finnish Road Administration
Traffic Engineering
Opastinsilta 12 A
P.O. Box 33
00521 HELSINKI
FINLAND
Telephone: Int. +358 204 2211

R. TAPIO LUTTINEN, RIKU NEVALA: Capacity and Level of Service of Finnish Signalized Intersections. Helsinki 2002, Finnish Road Administration, Finnra Reports 25/2002, ISBN 951-726-903-X, ISSN 1457-9871, TIEH 3200757E.

Keywords: capacity, level of service, traffic signal control

ABSTRACT

The current guidelines for the design of traffic signal control in Finland discuss very briefly the capacity and level-of-service issues. The procedures of the Swedish capacity manual from 1977 are still used. New Finnish and international research indicates that this method is outdated.

The base saturation flow rate for through vehicles according to the Swedish method and the Finnish guidelines is 1,700 pc/h. Recent Finnish research indicates a base value of 1,940 veh/h for a through lane. The saturation flow rate of turning movements without conflicts is 1,500 pc/h in the old guidelines. The new values are 1,800 pc/h for left turning movements and 1,750 pc/h for right and left+right turning traffic on a lane.

The results of three methods (American HCM2000, Danish DanKap, and Swedish Capcal 2) have been used to compare with the simulated control delays of the HUTSIM software calibrated to Finnish conditions. Two simple intersections, two example intersections from HCM, and one typical Finnish signalized intersection have been used as test cases. Both pretimed and traffic-responsive control were analyzed.

For pretimed control DanKap delay estimates were in a closer agreement with the simulation results than HCM2000. DanKap cannot, however, estimate the effect of traffic-responsive control. Because the difference in delay between an optimal pretimed control and a traffic-responsive control is small, and HCM and DanKap underestimate the delays at low and medium degrees of saturation, the DanKap delay estimates can be used also in the analysis of traffic-responsive systems. An optimized pretimed control should then be used for all the traffic conditions in the analyzed.

The interaction (stopped-time) delay estimates of Capcal 2 were closest to the simulated control delays. Of the three methods analyzed, Capcal 2 can be suggested as the best tool for the analysis of Finnish signalized intersections, both pretimed and traffic responsive. Capcal 2 does, however, have some convergence problems in the analysis of traffic-responsive control, and it is suggested that the cycle length and green splits are either entered manually, or at least checked for consistency.

There are no up-to-date guidelines for the capacity estimation of Finnish signalized intersections. The current manual presents both old and new saturation flow rates, but prefers the old values. It is obvious that this method underestimates the capacity of signalized facilities, and may suggest unnecessary investments. The application of adjustment factors for the effects of turning vehicles, opposing traffic and pedestrian conflicts would make the method more flexible and methodologically similar to the current international methods. In addition, the service measures and LOS criteria should be more clearly defined. There is an obvious need for new guidelines. This report can serve as a starting point for the update of the method.

R. TAPIO LUTTINEN, RIKU NEVALA: Capacity and Level of Service of Finnish Signalized Intersections [Valo-ohjauksisten liittymien välityskyky ja palvelutaso Suomessa]. Helsinki 2002, Tiehallinto, Finnra Reports 25/2002, ISBN 951-726-903-X, ISSN 1457-9871, TIEH 3200757E.

Asiasanat: välityskyky, palvelutaso, valo-ohjaus

TIIVISTELMÄ

Nykyinen valo-ohjauksen suunnittelun käsikirja (LIVASU 95) tarkastelee vain lyhyesti liittymien välityskykyä ja palvelutasoa. Käsikirja suosittelee edelleen vuoden 1977 ruotsalaiseen käsikirjaan perustuvaa menetelmää. Uudet suomalaiset ja kansainväliset tutkimukset osoittavat tämän menetelmän jo vanhentuneen.

Ruotsalaisen menetelmän ja LIVASU 95:n mukainen ominaisvälityskyvyn perusarvo on 1 700 ha/h. Uudempien suomalaisten mittausten antama perusarvo on 1 940 ha/h. Vanha perusarvo kääntyvän liikenteen ominaisvälityskyvyksi on 1 500 ha/h. Uudempien mittausten mukaiset ominaisvälityskyvyt ovat 1 800 ha/h vasemmalle ja 1 750 ha/h oikealle kääntyvien kaistalle sekä oikealle+vasemmalle kääntyvien kaistalle.

Kolmen laskentamenetelmän (amerikkalainen HCM2000, tanskalainen DanKap ja ruotsalainen Capcal 2) tuloksia verrattiin HUTSIM-ohjelmalla saatuihin keskimääräisiin ohjausviipeisiin. Testiliittyminä käytettiin kahta yksinkertaista perusliittymää, kahta HCM:n esimerkkiliittymää sekä yhtä LIVASU 95:n mukaista tyypillistä suomalaista valo-ohjattua liittymää. Laskelmissa tarkasteltiin sekä aikaohjausta että liikennetieto-ohjausta.

DanKapin viipeet olivat lähempänä HUTSIMin tuottamia viipeitä kuin HCM:n tulokset. DanKap ei kuitenkaan kykene arvioimaan liikennetieto-ohjauksen vaikutusta. Koska optimoidun aikaohjauksen ja liikennetieto-ohjauksen välisessä viivytyksessä stationaarisissa olosuhteissa on vain vähäinen ero, ja DanKap aliarvioi viipeet alhaisilla ja kohtuullisilla kuormitusasteilla, DanKapin viivearvioita voidaan käyttää myös liikennetieto-ohjatun liittymän analysoinnissa. Tällöin kaikissa liikennetilanteissa tulee tarkastella optimoitua aikaohjausta.

Capcal 2:n pysähtymisaika antoi parhaan arvion simuloiduista ohjausviipeistä. Kolmesta tarkastellusta menetelmästä Capcal 2:a voidaan pitää parhaana suomalaisten valo-ohjauksisten liittymien analysointiin, sekä aikaohjauksisten että liikennetieto-ohjauksisten. Capcal 2:lla oli kuitenkin joitakin konvergointiongelmia liikennetieto-ohjauksen analysoinnissa. Ohjelmaa käytettäessä onkin tarpeen syöttää kiertoaika ja vihreät ajat käsin, tai ainakin varmistaa ohjelman laskeman ajoituksen realistisuus.

Suomessa ei ole valo-ohjattujen liittymien välityskyvyn ja palvelutason arviointiin ajatasaista ohjetta. LIVASU 95 esittää sekä vanhat että uudet ominaisvälityskyvyn perusarvot, mutta suosittelee vanhojen arvojen käyttöä. On ilmeistä, että tämä menetelmä aliarvioi valo-ohjauksisten liittymien välityskyvyn ja saattaa johtaa ylimitoitettuihin investointeihin. Korjauskertoimien hyödyntämisen avulla nykyistä seitsemään kaistatyyppiin perustuvaa menetelmää voitaisiin kehittää joustavammaksi ja metodisesti paremmin nykyaikaisia menetelmiä vastaavaksi. Lisäksi palvelutason mittarit ja kriteerit tulisi määritellä nykyistä selkeämmin. Laskentaohjeiden tarkistamiselle onkin ilmeinen tarve. Tämä raportti tarjoaa lähtökohdan tarkistamistyölle.

PREFACE

The current guidelines for the design of traffic signal control in Finland discuss very briefly the capacity and level of service issues. The procedures of the Swedish capacity manual from 1977 are still used. New Finnish and international research indicates that this method is outdated.

New international analysis methods have been released in recent years, such as the American HCM2000, Danish DanKap, and Swedish Capcal 2. In order to apply these methods, it is important to know, how well they can describe the operational qualities of Finnish signalized intersections.

This report presents the results of a capacity and level of service research on Finnish signalized intersections. The results will be used in the development of capacity and level-of-service estimation methods in Finland.

The report was prepared by Dr. R. Tapio Luttinen from TL Consulting Engineers Ltd. The simulation experiments were conducted and reported by M.Sc. (Tech.) Riku Nevala, M.Sc (Tech.) Ville Lehmuskoski, and Dr. Jarkko Niittymäki from Helsinki University of Technology, Traffic and Transportation Laboratory.

The report is part of the Finnra strategic project S12 (Solutions to improve main roads). The work has been coordinated by deputy director Pauli Velhonoja at Finnra Traffic Engineering.

Helsinki 12th August 2002

Finnish Road Administration

Traffic Engineering

Contents

GLOSSARY

The terminology of the *Highway Capacity Manual* (Transportation Research Board 2000), and American terminology in general, is followed. Most of the definitions below are taken from the Highway Capacity Manual. Some adjustments have been made to make the definitions better applicable to the Finnish traffic signal control. The terminology for types of control is slightly modified. Phase-oriented definitions are expressed in terms of signal groups. The definitions for lost times follow the definitions in Exhibits 10-8 and 10-10 of HCM2000, which slightly differ from Exhibit 10-9.

Acceleration delay: Time lost due to the limited acceleration capability of a vehicle.

Actuation: Initiation or extension of a green interval through the operation of a detector.

Adjustment factor: A multiplicative factor that adjusts a capacity or service flow rate from one representing an ideal or base condition to one representing a prevailing condition.

Analysis period: A single time period during which a capacity analysis is performed.

Approach: A set of lanes at an intersection that accommodates all left-turn, through and right-turn movements from a given direction.

Base conditions: Characteristics for a given type of facility that are assumed to be the best possible from the point of view of capacity, that is, characteristics if further improved would not result in increased capacity. (Same as ideal conditions.)

Base saturation flow rate: The maximum steady flow rate—expressed in passenger cars per hour per lane—at which previously stopped passenger cars can cross the stop line of a signalized intersection under base conditions, assuming that the green signal is available and no lost times are experienced.

Bicycle: A vehicle with two wheels tandem, propelled by human power, and usually ridden by one person.

Calibration: The process of comparing model parameters with real-world data to ensure that the model realistically represents the traffic environment. The objective is to minimize the discrepancy between model results and measurements or observations.

Capacity: The maximum hourly rate at which persons or vehicles can reasonably be expected to traverse a point or uniform section of a lane or a roadway during a given time period under prevailing roadway, traffic, and control conditions.

Clearance lost time: The time interval at the end of the yellow change interval, which multiplied by the saturation flow rate gives the decrease of discharging vehicles because of the need to decelerate and stop.

Conflict groups: Signal groups that cannot give green signal indication simultaneously.

Conflicting movements: The traffic streams in conflict at an intersection.

Control condition: The traffic controls and regulations in effect for a segment of street or highway, including the type, phasing, and timing of traffic signals; stop signs; lane use and turn controls; and similar measures during the analysis period.

Control delay: The component of delay that results when a control causes vehicles on a lane or a lane group to reduce speed or to stop; it is measured by comparison with the uncontrolled condition.

Critical gap: The minimum time, in seconds, between successive major-stream vehicles, in which a minor-street vehicle can make a maneuver.

Critical lane group: The lane groups that have the highest flow ratio for a given signal phase.

Critical volume-to-capacity ratio: The proportion of available intersection capacity used by vehicles in critical lane groups.

Cycle: A complete sequence of signal indications.

Cycle length: The total time for a signal to complete one cycle.

Deceleration delay: Time lost due to deceleration.

Degree of saturation: Same as v/c ratio in traffic signals.

Delay: The additional travel time experienced by a driver, passenger or pedestrian.

Demand flow rate: The flow rate expected to desire service past a point or segment of the highway system at some future time, or the traffic currently arriving or desiring service past such a point.

Demand volume: The traffic volume expected to desire service past a point or segment of the highway system at some future time, or the traffic currently arriving or desiring service past such a point.

Demand to capacity ratio: The ratio of demand flow rate to capacity for a traffic facility.

Design application: Using capacity analysis procedures to determine the size (number of lanes) required for a specified level of service.

Detector: A device indicating a presence or passage of a vehicle or a pedestrian.

Deterministic model: A mathematical model that is not subject to randomness.

Discharge headway: The headway at stop line between two consecutive vehicles departing from a continuous queue of passenger cars.

Effective green time: The time during which the same number of vehicles can depart at a (constant) saturation flow rate as from a continuous queue during the actual green and yellow change intervals.

Effective red time: The cycle length minus the effective green time.

Equilibrium conditions: The expected state of the system may have cyclic fluctuations, but otherwise is time independent. (Steady state conditions)

Exclusive turn lane: A designated left- or right-turn lane or lanes used only by vehicles making those turns.

Flow rate: The equivalent hourly rate at which vehicles pass over a given point or section of a lane or roadway during a given time interval less than one hour. Also "rate of flow".

Flow ratio: The ratio of the demand flow rate to the saturation flow rate for a lane group at an intersection.

Fully actuated control: A signal control operation in which vehicle detectors at each approach to the intersection control the occurrence and length of every green interval.

Gap: The time, in seconds, for the front bumper of the second of two successive vehicles to reach the starting point of the front bumper of the first.

Gap acceptance: The process by which a minor-street vehicle accepts an available gap to maneuver.

Gap-seeking algorithm: An algorithm that uses detector information to extend the green time of a signal group until a large enough gap is found in the approaching traffic stream.

Geometric delay: The component of delay that results when geometric features cause vehicles to reduce their speed in negotiating a facility.

Green time: The duration, in seconds, of the green indication for a given movement at a signalized intersection.

Green time ratio: The ratio of the effective green of a signal group to the cycle length.

Headway: The time between successive vehicles as they pass a point on a lane or roadway, as measured from front bumper to front bumper.

Heavy vehicle: Any vehicle with more than four wheels touching the pavements during normal operation.

Ideal conditions: Characteristics for a given type of facility that are assumed to be the best possible from the point of view of capacity, that is, characteristics if further improved would not result in increased capacity. (Same as base conditions.)

Incremental delay: The second term of lane group control delay. It accounts for nonuniform arrivals and temporary random delays as well as delays caused by sustained periods of oversaturation.

Initial queue: The unmet demand at the beginning of an analysis period of a previous analysis period.

Initial queue delay: The third term of lane group control delay refers to the delays due to a residual queue identified in a previous analysis period and persisting at the start of the current analysis period.

Intergreen time: The time period between two consecutive green times in conflicting signal groups.

Interval: A period of time in which all traffic signal indications remain constant.

Lane group: A set of lanes established at an intersection approach for separate capacity and level-of-service analysis.

Lane group delay: The control delay for a given lane group.

Level of service: A qualitative measure describing operational conditions within a traffic stream, generally described in terms of such factors as speed and travel time, freedom to maneuver, traffic interruptions, comforts and convenience, and safety.

Lost time: The time during which and intersection is not used effectively by any movement. It is the sum of clearance lost time plus start-up lost time.

Major roadway: The roadway of a higher importance in an intersection.

Major flow: The traffic streams on a major roadway.

Maximum green: The maximum duration of a green interval. If there is green demand for conflict groups the green interval cannot extend beyond the maximum green. The starting point for the measurement of maximum green is given as a control parameter. Typical starting points are *i*) beginning of the green interval, *ii*) conflicting demand during the green, and *iii*) conflicting demand during the phase.

Measure of effectiveness: A quantitative parameter indicating the performance of a transportation facility or service.

Minimum green: The minimum duration of a green interval. Once started the green interval cannot end until after the duration of the minimum green.

Minor roadway: The roadway of a lower importance in an intersection.

Minor flow: The traffic streams on a minor roadway.

Operational application: A use of capacity analysis to determine the level of service on an existing or projected facility, with known or projected traffic, roadway, and control conditions

Opposing flow rate: The flow rate for the direction of travel opposite to the direction under analysis.

Overflow period: Time period during which demand exceeds capacity.

Passage detector: A detector which is able to detect the passage of a vehicle moving through the detection zone.

Passenger-car equivalent: The number of passenger cars having the same impedance effect as a single heavy vehicle of a given type, under prevailing roadway, traffic, and control conditions. (PCE)

Peak period: Time period during which arrival flow rate exceeds capacity.

Pedestrian: An individual traveling on foot.

Performance measure: A quantitative or qualitative characteristic describing the quality of service provided by a transportation facility or service.

Permitted plus protected: Compound turning movement protection that displays the permitted phase before the protected phase.

Permitted turn: A left or right turn at a signalized intersection that is made against an opposing or conflicting vehicular or pedestrian flow.

Phase: The part of the signal cycle allocated to any combination of traffic movements receiving the right-of-way simultaneously during one or more intervals.

Planning application: A use of capacity analysis to estimate the level of service, the volume than can be accommodated, or the number of lanes required, using estimates, HCM default values, and local default values as input.

Platoon: A group of vehicles or pedestrians traveling together as a group, either voluntarily or involuntarily because of signal control or other factors.

Presence detector: A detector which is able to detect the presence of a vehicle on the detection zone.

Pretimed control: A signal control in which the cycle length, phase plan, and phase times are preset to repeat continuously.

Prevailing condition: The geometric, traffic, and control conditions during the analysis period.

Progression adjustment factor: A factor used account for the effect of signal progression on traffic flow; applied only to uniform delay.

Protected plus permitted: Compound turning movement protection at a signalized intersection that displays the protected phase before the permitted phase.

Protected turn: A left or right turn at a signalized intersection that is made with no opposing or conflicting vehicular or pedestrian flow allowed.

Quality of service: A measure of the utilization of the transportation system.

Queue: A line of vehicles, bicycles or persons waiting to be served by a system.

Queue discharge: A flow with high density and low speed, in which queued vehicles start to disperse.

Queue discharge flow: A traffic flow of vehicles, bicycles or pedestrians queued during the red interval and crossing the stop line or curb line during the green interval.

Recreational vehicle: A heavy vehicle, generally operated by a private motorist, engaged in the transportation of recreational equipment or facilities; examples include campers, boat trailers, and motorcycle trailers.

Red clearance interval: An interval that follows a yellow-change interval and precedes a conflicting green interval in the next phase.

Red time: The period in the signal cycle during which, for a given lane group, the signal is red.

Red rest: Red signal indication during the rest state.

Rest state: The signal indication of a signal group after minimum green interval and green extensions, when there is no demand for the given signal group and its conflict groups.

Saturation flow rate: The hourly rate at which previously queued passenger cars can traverse an intersection approach under prevailing conditions, assuming that the green signal is available at all times and no lost times are experienced.

Saturation headway: The average time headway between vehicles occurring after the fourth vehicle in the queue and continuing until the last vehicle in the initial queue clears the intersection.

Semiactuated control: A signal control in which vehicle detectors at some approaches (typically on the minor street) control the occurrence and length of green intervals.

Service flow rate: The maximum hourly rate at which persons or vehicles can reasonably be expected to traverse a point or uniform section of a lane or roadway during a given time period under prevailing roadway, traffic, and control conditions while maintaining a designated level of service.

Service measure: A specific performance measure used to assign a level of service.

Service volume: The maximum hourly rate at which vehicles or persons reasonably can be expected to traverse a point or uniform segment of a roadway during an hour under specific assumed conditions while maintaining a designated level of service.

Signal group: The signal heads giving identical signal indications and controlling the same traffic movements.

Signalization condition: A timing diagram and/or the parameters of traffic-responsive control of a signalized intersection.

Simulation model: A computer program that uses mathematical models to conduct experiments with traffic events on a transportation facility or system over extended periods of time.

Spacing: The distance between two successive vehicles in a traffic lane measured from front bumper to front bumper.

Start-up lost time: The time, which multiplied by the saturation flow rate gives the decrease of discharging vehicles because of the need to react to the initiation of the green phase and to accelerate.

Stationary conditions: Conditions not changing during the observation period.

Steady state conditions: The expected state of the system may have cyclic fluctuations, but otherwise is time independent. (Equilibrium conditions)

Stochastic model: A mathematical model that employs random variables for at least one input parameter.

Stop time: A portion of control delay when vehicles are at complete stop. Also "stop delay".

Study period: A duration of time on which to base capacity analyses of a transportation facility.

Time-in-queue delay: Time spent in a queue; from stopping at the end of queue to passing the stop line.

Through vehicles: All vehicles passing directly through an intersection and not turning.

Traffic-actuated control: A signal control operation in which vehicle detectors at some or all approaches to the intersection control the occurrence and length of green intervals.

Traffic condition: A characteristic of traffic flow, including distribution of vehicle types in the traffic stream, directional distribution of traffic, lane use distribution of traffic, and type of driver population on a given facility during the analysis period.

Traffic-responsive control: A signal control operation in which gap-seeking or other algorithms use the information of vehicle detectors at each approach of the intersection to control the occurrence and length of green intervals.

Truck: A heavy vehicle engaged primarily in the transport of goods and materials or in the delivery of services other than public transport.

Undersaturation: A traffic condition in which the arrival flow rate is lower than the capacity.

Uniform delay: Delay at a signalized intersection assuming arrivals at uniform time headways.

Unit extension: The minimum time period between successive detector indications that will cause the signal controller to terminate the green display, assuming the absence of other demands for green extension.

Validation: Determining whether the selected model is appropriate for the given conditions and for the given task. It compares model prediction with measurements or observations.

Vehicle-actuated control: Same as traffic-actuated control. (VA control)

Volume to capacity ratio: The ratio of flow rate to capacity for a traffic facility. (v/c ratio)

Volume: The number of persons or vehicles passing a point on a lane, roadway, or other trafficway during some time interval.

Yellow change interval: The signal indication following the green interval.

SYMBOLS AND ABBREVIATIONS

A_i	Number of arrivals during the i^{th} cycle
$A(t)$	Number of arrivals during time $(0, t]$
$A'(t)$	Cumulative traffic demand during time $(0, t]$
C	Capacity
$C(\cdot)$	Control process
$C(t)$	Capacity curve
C_{i}	Capacity under ideal conditions
c	Cycle length
c_{max}	Maximum cycle length in traffic-responsive control
d	Distance
$d(t)$	Departure rate at time t
D_i	The i^{th} detector from the stop line
$D(\cdot)$	Demand estimator function (traffic model)
$D(t)$	Number of departures during time $(0, t]$
E_{B}	Equivalency factor for buses
E_{HV}	Equivalency factor for heavy vehicles
E_{R}	Equivalency factor for recreational vehicles
E_{T}	Equivalency factor for trucks
$\mathbb{E}[X]$	Expected value of random variable X
f_{a}	Adjustment factor for area type
f_{B}	Adjustment for buses
f_{bb}	Adjustment factor for stopping buses
f_{BC}	Adjustment factor for bicycles mixed with motor vehicles
f_{g}	Adjustment for grades
f_{HV}	Adjustment for heavy vehicles
f_{Lpb}	Pedestrian adjustment factor for left-turn movements
f_{LT}	Adjustment factor for left turns in the lane group
f_{LU}	Adjustment factor for lane utilization
f_{P}	Adjustment factor for the effect of parking
f_{p}	Uniform delay progression adjustment factor
f_{R}	Adjustment for recreational vehicles
f_{Rpb}	Pedestrian adjustment factor for right-turn movements
f_{RT}	Adjustment factor for right turns in the lane group
f_{s}	Adjustment for platoon arrivals during green
f_{T}	Adjustment for trucks
f_{va}	Adjustment factor for vehicle-actuated control
f_{w}	Adjustment factor for roadway width
$f(x)$	Probability density function
$F(x)$	Probability distribution function
Finnra	Finnish Road Administration
g	Length of a green interval
g_{e}	Effective green
g_{max}	Maximum effective green length
g_{min}	Minimum effective green length
h_{d}	Average discharge headway
h_i	Average discharge headway of movement i
HCM	*Highway Capacity Manual*
HUT	Helsinki University of Technology
I	Variance-to-mean ratio, upstream filtering/metering adjustment factor

I_A	Variance-to-mean ratio of a arrival counting process
I_D	Variance-to-mean ratio of a departure counting process
k	Incremental delay factor
L	Total lost time (time interval between effective greens)
$L(\cdot)$	Control logic function
$L(t)$	Queue length at time t
L_i	Queue length at the end of the i^{th} cycle
$L_o(t)$	Overflow queue at time t
L_v	Vehicle length
l	Lost time ($l_1 + l_2$)
l_1	Start-up lost time
l_2	Clearance lost time
LOS	Level of service
$M(\cdot)$	Traffic measurement function (detection)
M/D/1	Queuing system with random arrivals and one server with deterministic service times
M/G/1	Queuing system with random arrivals and one server with a general service time distribution
$\min\{x_1, \dots\}$	Minimum element in a set
MOE	Measure of effectiveness
N	Number of lanes in the lane group
$\mathbb{N}$	The set of natural numbers $\{1, 2, 3, \dots\}$
N_A	Number of approaches in an intersection
N_P	Number of phases
$N_S(t)$	Number of stops during time $(0, t]$
P_B	Proportion or percentage of buses in traffic flow
P_g	Proportion of green arrivals
P_{HV}	Proportion or percentage of heavy vehicles in traffic flow
P_{LR}	Proportion or percentage of turning vehicles on a shared lane
P_{LT}	Proportion of left turning vehicles
P_R	Proportion or percentage of recreational vehicles in traffic flow
P_r	Probability of all-red during a no-demand condition
P_{RT}	Proportion of right turning vehicles
P_T	Proportion or percentage of trucks in traffic flow
p_i	Proportion of movement i in a lane group
pc	Passenger car
PCE	passenger-car equivalent
pcu	Passenger car unit
q	Flow rate, traffic volume
$\hat{q}_k$	Traffic demand estimate for time interval k
q_o	Opposing flow rate
q_p	Overflow peak flow rate
q_{ped}	Pedestrian flow rate
$q(t)$	Flow rate at time t
Q_i	Service flow rate for level of service i
R	Radius
r	Length of a red interval
R^2	Coefficient of determination
s	Saturation flow rate
s_0	Base saturation flow rate

s_i	Saturation flow rate for lane group i
s_k	Saturation flow rate for time period k
s_B, s_E ...	Saturation flow rate for lane type B, E...
s_{LT}	Basic saturation flow rate for permitted left turn movements
t	Time
t_0	Time of queue discharge
t_a	Arrival time from stop line to point of conflict
t_b	Blocked time
t_d	Departure time from stop line to point of conflict
t_o	Length of overflow period
t_p	Length of peak period
t_y	Length of yellow change interval
u_k	The state of signal control during time interval k
v	Speed
v_L	Speed limit
VA	Vehicle actuated
v/c	Volume-to-capacity ratio
W	Delay
W_a	Acceleration delay
W_d	Deceleration delay
$W_o(t)$	Total overflow delay during time $(0, t]$
W_s	Stop delay
$W_u(t)$	Total uniform delay during time $(0, t]$
w	Average control delay
w_0	Average delay under very low flow conditions
w_a	Average acceleration delay
w_D	Average delay in an M/D/1 queuing system, random delay component of average control delay
w'_D	Deterministic delay component with maximum green and cycle lengths
w_d	Average deceleration delay
w_G	Average delay in an M/G/1 queuing system
w_i	Delay of vehicle i
w_l	Lane width
$w_o(t)$	Average overflow delay during time $(0, t]$
$w'_o(t)$	Average overflow delay per arriving vehicle during time $(0, t]$
w_q	Initial queue delay component of average control delay
w_r	Time dependent random delay component of average control delay
w_s	Average stop delay, shoulder width
w_u	Average uniform delay
X_+	Positive part of X; i.e., $X_+ = (X + \|X\|)/2$
x_a	Acceleration distance
x_d	Deceleration (braking) distance
Y	Sum of critical flow ratios in the cycle
y	Flow ratio (q/s)
y_j	Critical flow ratio of phase j
z_k	The state of traffic during time interval k
β	Deceleration rate
γ	Grade (percent)
Δ	Extension time
δ	Time step length in a discrete time model

λ	Average arrival rate
μ_k	The state of detectors during time interval k
$\mu_S(i)$	Membership function; 1 if $i \in S$, 0 otherwise
ρ	Degree of saturation, utilization factor
ρ^*	Degree of saturation with maximum effective green and maximum cycle length
ρ'	Utilization factor
$\rho_{\max}$	Maximum degree of saturation for any lane
ρ_p	Degree of saturation during the peak flow rate
σ	Standard deviation
σ^2	Variance
τ	Average service time

1 INTRODUCTION

Rouphail, Anwar, Fambro, Sloup & Perez (1997) used simulations and field measurements to develop the delay model of the Signalized Intersections chapter in the 1994 HCM (Transportation Research Board 1994) more suitable for traffic-actuated control. The "generalized delay model" included a new parameter k derived from queuing theory. It described the properties of departure distributions at the stop line. The generalized delay model produced traffic-actuated delay values comparable to field measurement and simulations. The model was also shown to be sensitive to volume changes and traffic-actuated control settings. This model was introduced into the 1997 update of HCM (Transportation Research Board 1998), and with slight modifications into HCM2000 (Transportation Research Board 2000). The service measure of signalized intersections in HCM2000 is control delay.

HCM provides the most wide-spread method for the planning, design, and operational analysis of transportation facilities. It is calibrated for North-American conditions. There are, however, significant differences in the driver behavior, types of vehicles, and the operation of traffic signals between North America and the Nordic countries. Consequently, it is important to compare the HCM method with the Nordic methods, most important of which are the Danish DanKap (Vejdirektoratet 1999*b*) and the Swedish Capcal 2 (Swedish National Road Administration 1995).

In Finland, the planning and analysis of traffic signal control has been based on the old Swedish method (Statens vägverk 1977), which is an application of the Webster method (Webster 1958, Webster & Cobbe 1966). New Finnish and international research indicates that the current Finnish procedure is now outdated. Also, there is no delay evaluation method designed especially for Finnish traffic conditions and traffic-responsive signal control, neither have the current international methods been tested under Finnish conditions.

The object of this study is to evaluate how well the procedures of HCM2000, DanKap, and Capcal 2 are able to predict the delays in Finnish intersections controlled by pretimed or traffic-responsive signals. The delays calculated with these methods have been compared with the delay results of HUTSIM simulations. The comparisons have been carried out in five test intersections to obtain an overall picture of the suitability of these methods under Finnish conditions.

A theoretical overview of traffic signal control is presented in the first chapters to describe the central concepts. A description of the Finnish traffic signal control philosophy and procedures in section 3.2.3 is both informative for international readers, and highlights issues which may explain differences between Finnish and international results. Considerable attention has been given to the theory of delay estimation in Chapter 4. The discussion gives a theoretical overview of the subject, and helps to understand the delay estimation methods in the international procedures. The description of the Finnish saturation flow rates in section 5.4 is based on the HUTSIM calibration studies by the Helsinki University of technology. No new measurements have been performed for this research.

The results of HCM2000, DanKap, Capcal, and the Webster method have been compared with the control delays obtained with the HUTSIM simulation software. The parameters of this software have been calibrated for Finnish conditions. Five test intersections have been analyzed with both pretimed and traffic-responsive control.

Only isolated intersections are discussed. HCM2000 provides methods for the analysis of urban streets (Chapter 15) as well as corridors (Chapter 29) and wider areas (Chapter

30). These facilities are beyond the scope of the current research. The pedestrian level of service has been only briefly discussed. The estimation of pedestrian delay has been considered a trivial effort and has not been much attention in this research. The effects of pedestrians, heavy vehicles and intersection geometry (lane width, turning radius, grade) on vehicular delays have not been studied because of the limitations in the simulation methodology.

In many signalized intersections priorities are arranged for public transportation (buses and trams). In these cases the operational analysis should give different weighing factors to different vehicle categories (Vincent 1973). This important topic is, however, beyond the scope of the current research.

Other important issues in recent discussion include the usefulness of the level-of-service classification (Kittelson & Roess 2001), user perception of service quality (Pécheux, Pietrucha & Jovanis 2000) and the uncertainties in the analysis results (Heydecker 1987, Khatib & Kyte 2001). Mäkelä (1997) has presented a wider framework for the level-of-service criteria of Finnish signalized intersections. At this time it was considered appropriate to wait for the conclusions of the international discussions.

2 TRAFFIC SIGNAL CONTROL

2.1 Principles of traffic signal control

Traffic signal control separates traffic streams in an intersection by allocating different time intervals for conflicting traffic movements. For each movement the signals are given cyclically in the following order:

1. Red
2. Red + yellow
3. Green
4. Yellow

Green signal indicates "go" and red indicates "stop". A green interval is followed by a yellow change interval indicating that a vehicle must stop if it can be done safely. The length of the yellow change interval depends on the speed limit according to table 2.1. A red+yellow signal indicates that the green signal will be given shortly. The length of the red+yellow interval is 1.0–1.5 seconds.

Table 2.1: The length of the yellow change interval

Speed limit (km/h)	Yellow (s)
70	5
60	4
40–50	3

Because different traffic movements may use same lanes, and movements on separate lanes may be controlled by the same signal indications, the physical unit of signal control is a *lane group*. It consists of one or more lanes in an approach.

The signal indications may be either circular or arrow shaped. A green arrow indicates that vehicles may execute the movement in the direction of the arrow. The turning movement controlled by an arrow is called "protected"; i.e., it is free of conflicting vehicular or pedestrian traffic. A circular green indicates that vehicles may proceed cautiously. This movement is protected from primary conflicts, but secondary conflicts are possible. *Primary conflicts* occur between movements originating from rectangular approaches (Fig. 2.1). Other conflicts are *secondary conflicts*. A turning movement with secondary conflicts allowed is called "permitted".

Each lane group is controlled by a *signal group* consisting of a primary signal (typically beside or above the stop line) and one or more secondary signals. One of the secondary signals must be at least 2.5 meters behind the primary signal. The signal heads of a pedestrian signal group are at the both ends of the crosswalk or separately controlled parts of the crosswalk.

A signal head has typically three lenses (red, yellow and green). It is, however, possible to use four or five-lens arrangements according to Figure 2.2 to indicate different combinations of protected and permitted turning movements. The protected-plus-permitted left turn (Fig. 2.3) has been considered hazardous, and it is not allowed in Finland (Liikenne- ja viestintäministeriö 2000). A permitted-plus-protected left turn is possible.

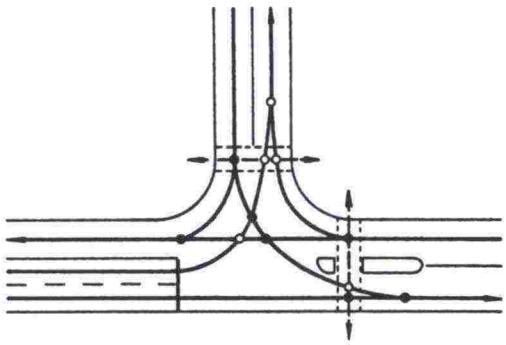


Figure 2.1: Primary (filled circles) and secondary (empty circles) conflict points in a T-intersection (Pohjoismaiden tieteknillinen liitto 1978)

Right turn on red (RTOR) is not allowed in Finland. It is, however, possible to have a triangular island that separates right turning vehicles from the signalized traffic flows. These vehicle have to observe a yield sign.

Green indications of conflicting movements are separated by *intergreen times*. They ensure the safe and logical operation of traffic signals. An integreen time cannot be shorter than yellow change interval plus red clearance interval, which give time for vehicles to pass all conflict points (Fig. 2.1) before the arrival of the vehicles in the next phase. The intergreen time between signal groups i and j is measured from the end of a green interval of signal group i to the beginning of the green interval of signal group j (Fig. 2.4).

Assuming that vehicles at the end of green i depart at speed v_i and vehicles at the beginning of green j approach the conflict point at speed v_j the safety interval (minimum intergreen time required for safety) is

$$S_{i,j} = t_y + t_d - t_a, \tag{2.1}$$

where t_y is the yellow change interval, t_d is the departure time, and t_a is the arrival time. The departure time is

$$t_d = \frac{d_i + L_v}{v_i}, \tag{2.2}$$

where d_i is the distance of the departing vehicle from the stop line to the conflict point and L_v is the length of the vehicle. The arrival time is

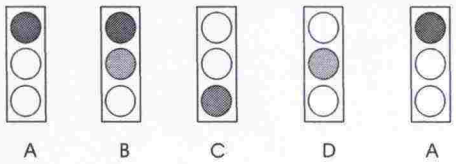
$$t_a = \frac{d_j}{v_j}, \tag{2.3}$$

where d_j is the distance of the arriving vehicle from the stop line to the conflict point. The minimum intergreen time between signal groups i and j is the largest safety interval between all feasible conflict points.

Intergreen times are assigned between all signal groups that cannot display green indications simultaneously. Intergreen time may be larger than the maximum safety interval in order to achieve a proper synchronization between the start times of green intervals in the same phase. It is also possible to set intergreen times for non-conflicting movements. A typical application of a "virtual intergreen time" is the synchronization of green intervals at a crosswalk controlled by two signal groups.

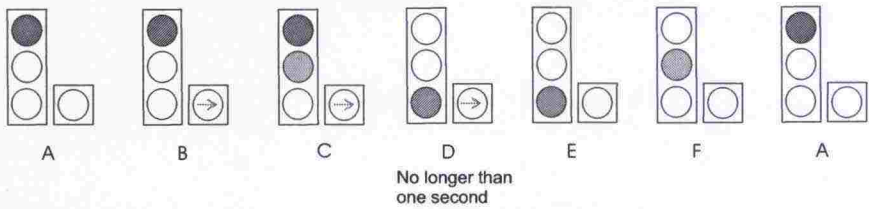
The part of the signal cycle allocated to any combination of traffic movements receiving the right-of-way simultaneously during one or more intervals is called a *phase*. The

Three-lense arrangement

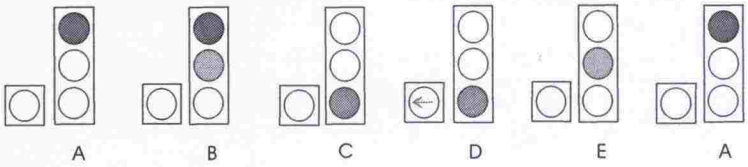


Four-lense arrangements

Protected plus permitted right turn

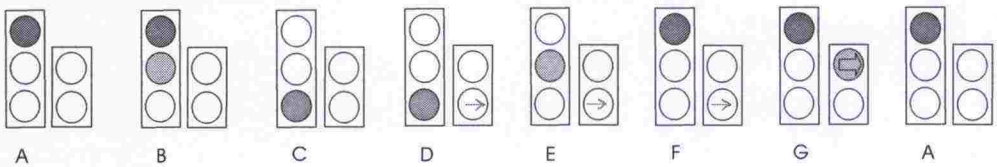


Permitted plus protected left turn

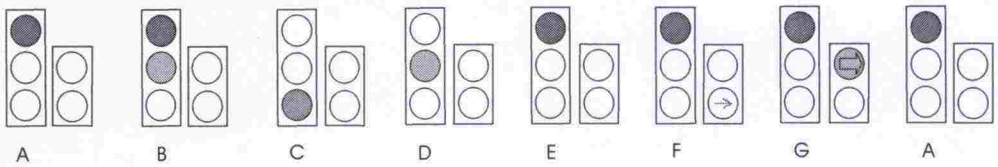


Five-lense arrangements

Permitted plus protected right turn



Additional right-turn phase



Five-lense arrangement at railroad crossings

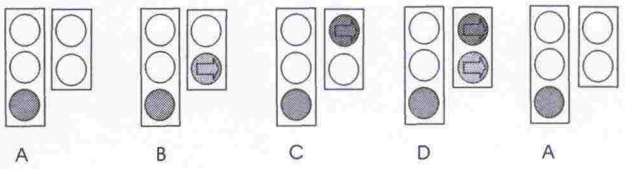


Figure 2.2: Allowed lense arrangements in Finland (Kehittämiskeskus 1996)

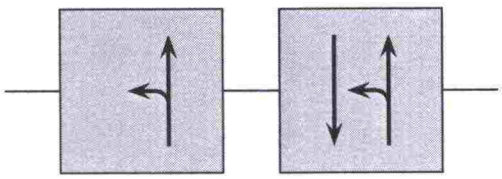


Figure 2.3: Protected-plus-permitted left turn (prohibited in Finland)

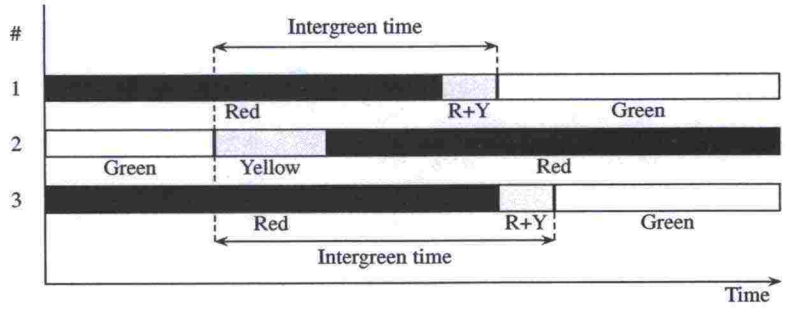


Figure 2.4: Timing plan for signal groups

start and end times of green intervals between signal groups in the same phase may not be simultaneous, but they depend on safety intervals and other possible considerations.

2.2 Effective green

As a vehicle approaches an intersection displaying a red signal the driver decelerates and stops either at the stop line or at the end of a queue (fig. 2.5). When the signal turns green the driver accelerates until the vehicle reaches its desired or maximum possible speed.

The red+yellow signal indicates that the green phase is about to begin. The drivers of the first vehicles become alert and prepare to start moving. As the green begins the first vehicle starts to accelerate. The red+yellow signal helps the first driver to anticipate the starting green interval (see Fig. 2.6)

The discharge process of the vehicles in the queue is controlled by the reaction times and desired acceleration rates of drivers as well as the acceleration rates of the vehicles ahead. At the beginning of the green interval the discharge rate at the stop lane starts to increase. As the queuing vehicles have reached a constant speed at the stop line the discharge rate has reached its maximum, called the *saturation flow rate*. On average, the discharge headways reach a constant level of slightly below two seconds after the fourth vehicle. The saturation flow rate may vary from cycle to cycle, but an average value can be used for given conditions.

Figure 2.6 displays a comparison of Finnish discharge headways (Niittymäki & Pursula 1997) with the results of Briggs (1977) and Teply & Jones (1991). King & Wilkinson (1976) presents results of older discharge headway studies, which do not reach as low values as the later results in figure 2.6.

As the green interval ends the approaching drivers make a decision whether to continue across the stop line or stop. At a macroscopic level the departure rate (under saturated conditions) starts do decrease and reaches zero as the red phase begins, or soon after if.

Figure 2.7 displays a departure rate curve $d(t)$ under saturated conditions. The number

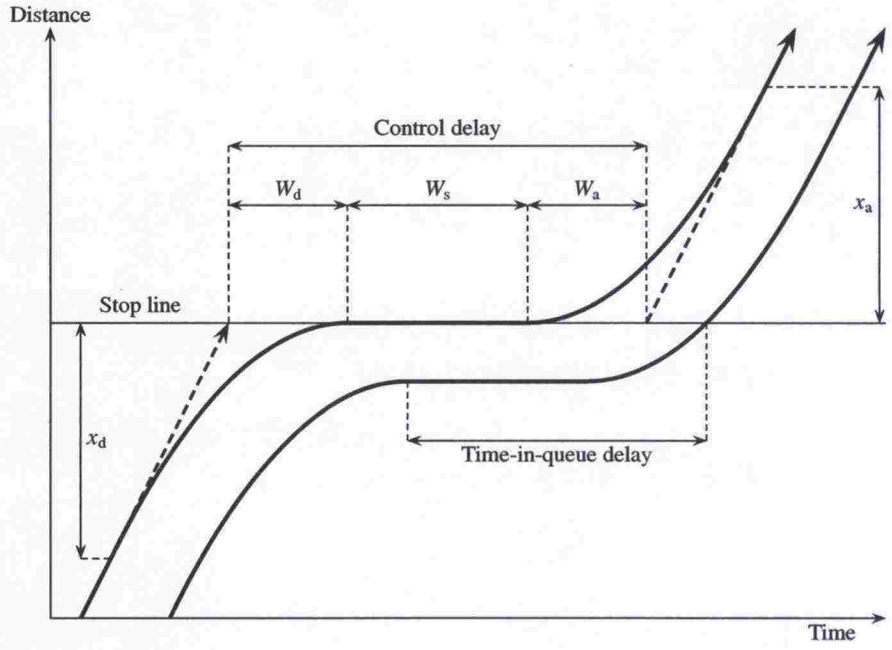


Figure 2.5: Delays at traffic signals

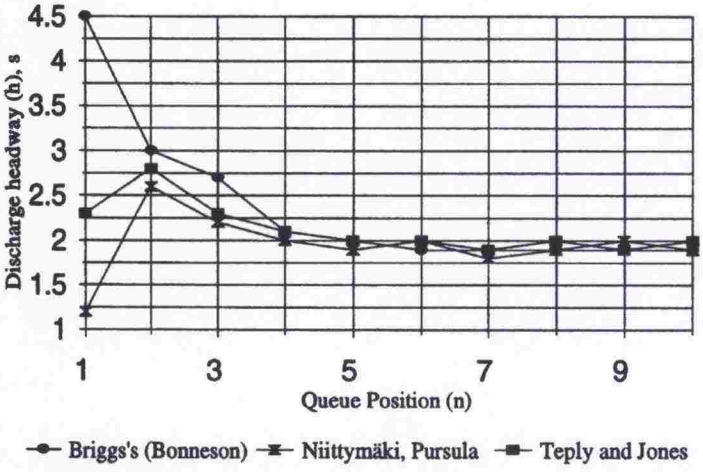


Figure 2.6: Discharge headways as a function of queue position (Niittymäki & Pursula 1997)

of departures during a cycle length (c) is equal to the area under the departure rate curve:

$$D(c) = \int_0^c d(t)\,\mathrm{d}t. \tag{2.4}$$

The *effective green* (g_e) is the green time required for $D(c)$ departures assuming that the departure rate during the effective green is constant and equal to the saturation flow rate (s):

$$D(c) = g_e s \tag{2.5}$$

(Webster 1958). That is, the start-up lost time (l_1) and the clearance lost time (l_2) are compensated, and the area $D(c)$ is equal to the area of the gray rectangle $g_e s$ in figure 2.7.

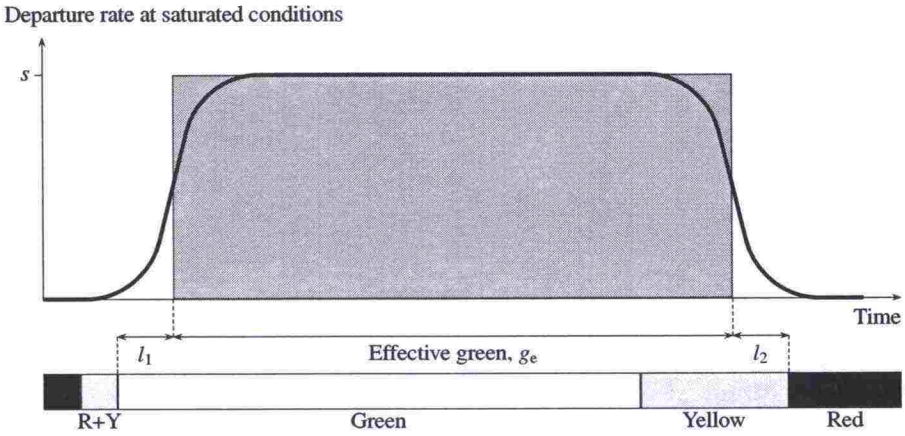


Figure 2.7: Departure rates and effective green

It is usually (Clayton 1941, Webster 1958) assumed that after startup lost time the saturation flow rate remains constant until the beginning of the yellow change interval. The effective green time is

$$g_e = g + y - l = g + y - l_1 - l_2, \tag{2.6}$$

where the lost time (l) is the sum of startup lost time (l_1) and clearance lost time (l_2). In Finland $l = 0.5t_y$ is used in the cycle length calculations (Kehittämiskeskus 1996).

In the discussion below on green time (g) is assumed to be equal to the effective green (g_e).

2.3 Option zone and dilemma zone

Let us assume that a vehicle is approaching an intersection at speed v. At a distance d from the stop line the driver starts to decelerate at a constant rate β so that the vehicle stops at the stop line. The deceleration rate is

$$\beta = -\frac{v^2}{2d}. \tag{2.7}$$

resulting in the deceleration time

$$t_d = \frac{2d}{v}. \tag{2.8}$$

The deceleration delay is the time difference between the deceleration time and the time to traverse distance d at speed v:

$$W_d = \frac{2d}{v} - \frac{d}{v} = \frac{d}{v}. \tag{2.9}$$

For a maximum accepted deceleration rate of β_{max} the minimum stopping distance is

$$d_{min} = vt_r + \frac{v^2}{2\beta_{max}}, \tag{2.10}$$

where t_r is the perception time plus reaction time of the driver. If the distance of a vehicle from the stop line is shorter than d_{min}, the vehicle cannot stop before the stop

line. When the signal turns yellow a vehicle can cross the stop line during the yellow change interval (t_y), if the distance from the stop line is not longer than

$$d_y = vt_y. \tag{2.11}$$

If $d_y < d_{min}$ and the vehicle is in the zone $d_y \ldots d_{min}$, it can neither stop before the stop line nor cross the stop line before the signal turns red. This zone is called the *dilemma zone*. If $d_{min} < d_y$, a vehicle in the zone $x_s \ldots x_g$ can either stop before the stop line or pass the stop line during the yellow change interval. This zone is called the *option zone*.

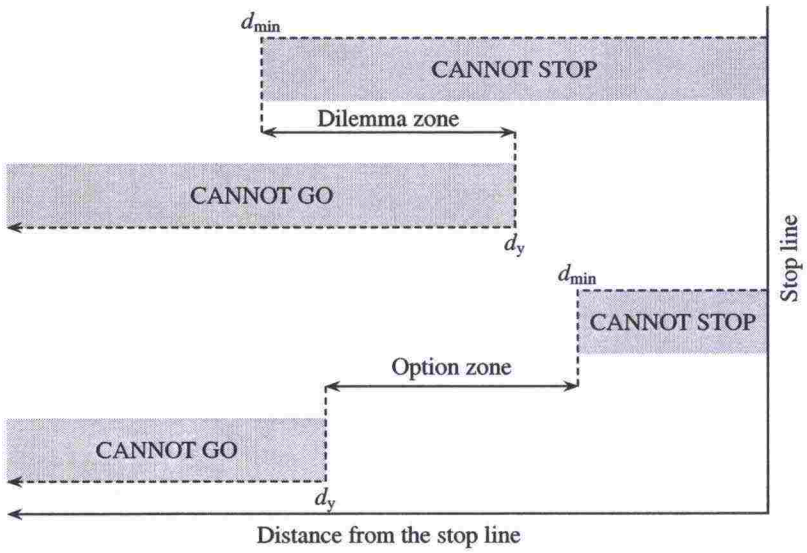


Figure 2.8: Dilemma and option zones

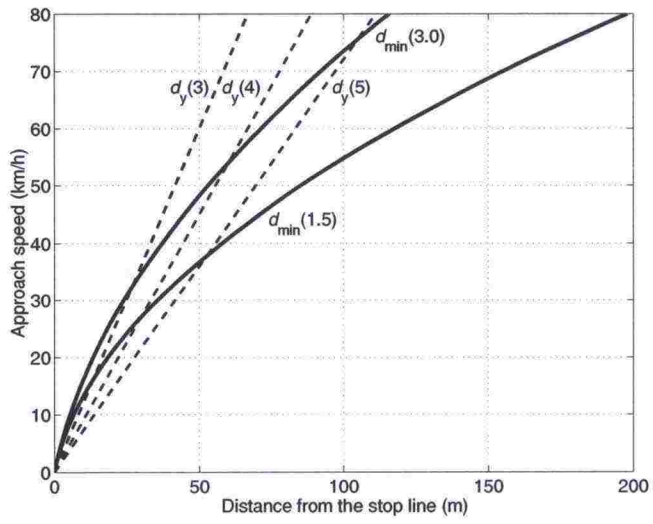


Figure 2.9: Maximum go-distances (d_y, dashed lines) for yellow change intervals of 3, 4, and 5 s, and minimum stop-distances (d_{min}, solid lines) for deceleration rates of 1.5 and 3.0 m/s^2

2.4 Vehicle dynamics at signalized intersections

HUTSIM calibration studies at four Finnish intersections indicated that the average acceleration of passenger cars was in the range 1.55–2.14 m/s^2. The average deceleration of passenger cars varied in the range 1.55–2.20 m/s^2. (Niittymäki & Pursula 1994).

The acceleration model installed in HUTSIM is speed dependent. Acceleration varies from 2.6 m/s^2 to 0.7 m/s^2 as a function of speed (Niittymäki 1998).

For other vehicle classes the average acceleration and deceleration are given in tables 2.2 and 2.3. The average acceleration and deceleration rates for buses are approximately 1.2 m/s^2. The acceleration and deceleration rates of trams are 1.2 m/s^2 and 1.3 m/s^2 (Niittymäki 1998).

Table 2.2: Acceleration (m/s^2) at Finnish signalized intersections (Niittymäki & Pursula 1994)

Vehicle type	Average m/s^2	Standard deviation
Van	1.50	0.37
Lorry	1.20	0.07
Semitrailer	1.09	0.11
Full trailer	1.05	0.12

Table 2.3: Deceleration (m/s^2) at Finnish signalized intersections (Niittymäki & Pursula 1994)

Vehicle type	Average m/s^2	Standard deviation
Van	1.82	0.48
Lorry	1.69	0.31
Semitrailer	1.66	0.42
Full trailer	1.43	0.27

The average speed of pedestrians on signalized crosswalks is approximately 1.30–1.45 m/s. The 15-percentile speed is approximately 1.2 m/s, which is the speed used in planning applications (Niittymäki 1998). In HUTSIM the average pedestrian speed is 1.4 m/s (Niittymäki & Pursula 1994).

The basic time gap between trailing vehicles in HUTSIM is 1.3 s. The base value for critical gap in permitted left turns is 4.0 s. (Niittymäki 1993)

3 CONTROL METHODS

3.1 Pretimed control

The first traffic signal was installed in Westminster, U.K. It was of the semaphore-arm type with red and green gas lamps for night use. The experiment ended in an explosion. The first three-colored traffic signals were installed in New York in 1918. They were operated manually. (Webster & Cobbe 1966) Current traffic signals operate automatically using either pretimed or traffic-responsive control.

Pretimed control gives the signal indications using a fixed timing plan. The timing plan, or control logic, may change according to a given schedule. The state of traffic (arriving, queuing and discharging vehicles) do not affect the control (Fig. 3.1).

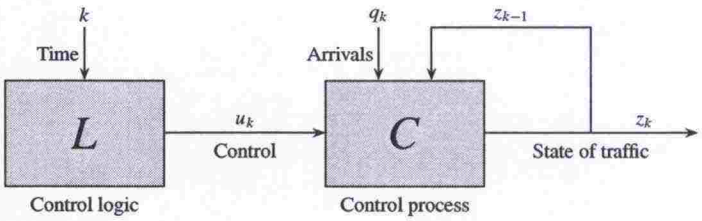


Figure 3.1: Pretimed traffic signal control process (Luttinen 1988)

The state (z_k) of the traffic at a very short time interval k is determined by the state (z_{k-1}) of the traffic in the preceding time interval, and the arrivals (q_k) and the signal control (u_k) during the observed time interval k: i.e.,

$$z_k = C(z_{k-1}, q_k, u_k). \tag{3.1}$$

In a pretimed system the control (u_k) is determined by time (k) only:

$$u_k = L(k). \tag{3.2}$$

Webster (1958) and Webster & Cobbe (1966) have presented the most widely used method for setting cycle lengths and green intervals for pretimed signals at undersaturated conditions. The earliest studies of pretimed control at oversaturated conditions were presented by Gazis (1964) and Gazis & Potts (1965). At an isolated oversaturated intersection the settings should be such that after the peak period the queues at critical directions are cleared simultaneously.

3.2 Traffic-responsive control

3.2.1 Types of controllers

The first traffic-actuated signals were installed at the beginning of the 1930s (Watson 1933, Webster & Cobbe 1966). Early models of traffic-actuated controllers were electromechanical. Later controllers were of solid-state design (Fig. 3.2). Since early 80's the controllers have been based on microprocessors with an extensive number of parameters and programmability.

In a *traffic-actuated* system the control during time interval k is based on the state of the control and detector pulses in the preceding time interval $k-1$ (Fig. 3.3). Time has two functions: The parameters of the control logic (such as unit extension) are time

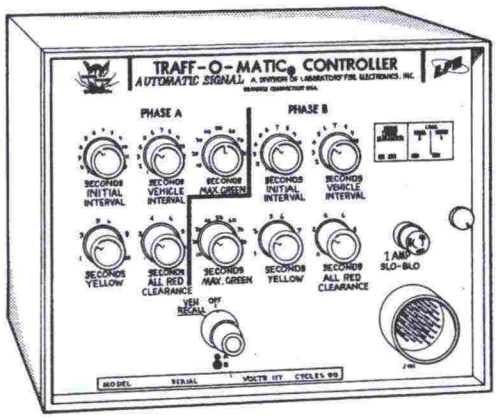


Figure 3.2: A solid-state, two-phase, fully actuated controller in the early 1970's (Pignataro 1973)

dependent, and the parameters (such as maximum green) may change according to a time schedule. The signal control can be expressed as

$$u_k = L(u_{k-1}, \mu_{k-1}, k). \tag{3.3}$$

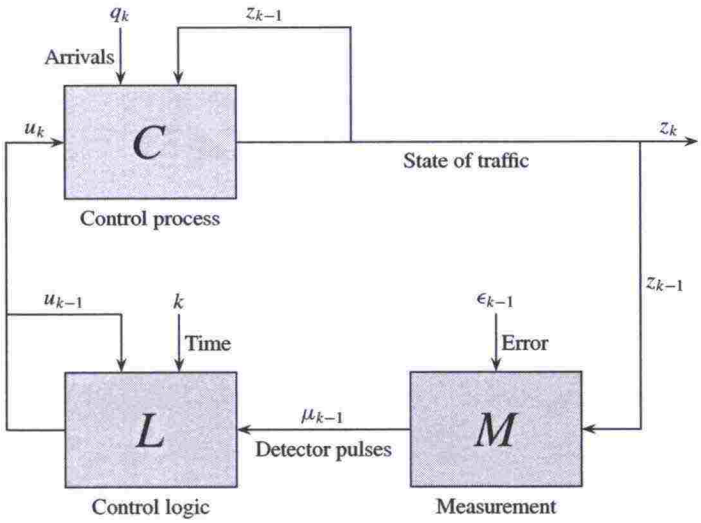


Figure 3.3: Traffic-actuated traffic signal control process (Luttinen 1988)

Traffic-actuated controllers have been classified as fully actuated, semiactuated, and volume-density controllers. The properties of early two-phase *fully actuated controllers* have been described by Hammond & Sorenson (1941):

1. Right-of-way is not given to any street without actuation thereon, and, in the complete absence of traffic, will remain upon the street where it was last assigned; an exception being that recall switches are provided which when thrown will cause the right-of-way to revert to a selected street despite actuation thereon.

2. The actuation of a detector on one street while right-of-way is on the other street, causes right-of-way to be transferred to the first street, only after a minimum adjustable interval has elapsed and after proper intergreen time, as follows:

(a) Immediately, providing there has been no actuation on the other street for a definite adjustable interval.

(b) After a predetermined maximum interval despite continued actuation on street having right-of-way.

(c) After an adjustable prefix interval has expired, each detector actuation causes the controller to cancel the time extension interval then in effect and begin the timing of another extension interval.

(d) If the maximum interval takes the right-of-way from a street before the last detector actuation has allowed the vehicle to enter the intersection, or if detector actuation occurs during the red or the clearance interval on any street, right-of-way is returned to the street at first opportunity and without the necessity of further actuation.

Although at the beginning of 1940's a slightly different terminology was used, the key parameters of traffic-actuated control were the demand (actuation), unit extension interval, minimum green interval, and maximum green interval. See Pignataro (1973) for a later description.

A *semiactuated controller* uses detector information to transfer the right-of-way to the side street upon traffic demand. The green may extend up to a predetermined maximum, after which the right-of-way returns to the major street. The green interval of the major street has a preset minimum duration. (Hammond & Sorenson 1941)

Volume-density controllers were advanced fully actuated controllers, which were able to take into account traffic volumes, densities, and elapsed waiting time on each traffic phase (Evans 1950). Detectors were located farther back from the stop line than in the case of the (then) conventional fully actuated control. Pignataro (1973) has described the essential operating features of this controller type as follows:

1. Detectors on all approaches are placed sufficiently back from the intersection to enable the counting of relatively large numbers of queued vehicles.

2. Each phase has an assured green time, as set by three dials:

   (a) Minimum green

   (b) Number of actuations before minimum green is increased

   (c) Extension of minimum green for each additional actuation during the red phase

3. Passage time is the unit extension, created by each additional actuation, after the assured green time has elapsed. Time is set for a vehicle to travel from the detector to the stop line. This value also becomes the maximum allowable gap between vehicles which will retain the green. This maximum gap, or passage time, may be reduced in several ways. The green will be lost when:

   (a) A pre-determined low limit of passage time is reached when red-phase vehicles have waited a preset time.

   (b) A pre-determined low limit of passage time is reached when the number of vehicles waiting on the red phase exceeds a preset value.

   (c) A pre-determined low limit of passage time is reached when the number of green-phase vehicles per ten seconds is less than a preset value.

These constitute the density function of the controller. Special dial settings differ from model to model.

4. Platoon carryover effect enables the controller to remember a preset percentage of the previous green-period traffic and synthetically applies that number of vehicles waiting on the red phase, when the next platoon of vehicles hits the detector.

5. Green extension is limited to a preset extension limit. However, this feature seldom operates, because of the effect of the reduction factors on passage time.

6. Clearance intervals are preset for each phase.

7. Each phase has a recall switch that operates in the same manner as for the fully actuated controller:

   (a) With all recall switches off, the green indication remains on the phase to which it was last called, provided there is no actuation on the other approaches.

   (b) With a recall switch on, the green indication will revert to that selected phase at every opportunity.

   (c) With the recall switch on both phases of a two-phase control, the controller operates on a pretimed basis, provided there is no demand on either phase.

See also Kell & Fullerton (1982). Orcutt (1993) discusses modern volume-density control under the name gap-reduction control.

The features of current microprocessor-based controllers far exceed the old volume-density controllers, and the distinction between fully actuated and volume-density controllers is currently obsolete. In HCM2000 this type of control is called traffic actuated. In this report "traffic-actuated signals" describe an implementation of a gap-seeking algorithm. In Britain this method is referred to as System D (Heydecker 1990). "Traffic-responsive control" is preferred as a more general description of a closed-loop or adaptive control, including the control strategies presented below.

3.2.2 Control strategies

The traditional traffic-actuated controllers extended the green phase until a gap larger than a unit extension interval was found in the arriving traffic flow. The volume-density controllers extended this *gap-seeking algorithm* as described above. Traffic-actuated control has usually been applied on isolated intersections, but the gap-seeking method can also be used to adjust coordinated control to local traffic demand (Jovanis & Gregor 1986).

Morris & Pak-Poy (1967) observed that traffic-actuated signals based on a gap-seeking algorithm increased capacity and decreased the average delay to one half of the average delay at an equivalent pretimed signal. They observed that the optimum (delay minimizing) unit extension interval decreased as flow rate increased (Fig. 3.4). Maximum and minimum green functions decreased the efficiency and capacity of signal control.

An obvious alternative to gap-seeking algorithms is *demand-responsive control* (Fig.3.5), which adjusts the control to estimated demand on each approach. A traffic model (D) uses detector and control information to update the current demand estimate:

$$d_k = D(d_{k-1}, u_{k-1}, \mu_{k-1}). \tag{3.4}$$

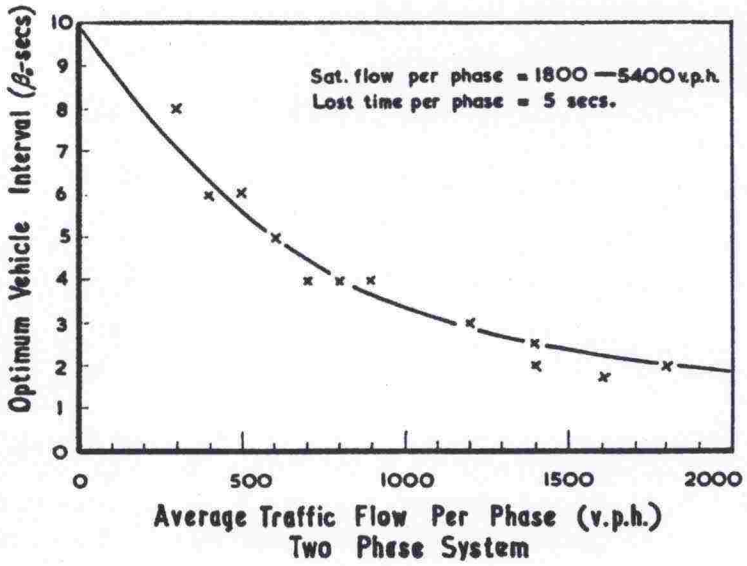


Figure 3.4: Variation in optimum unit extension interval with average traffic flow per phase at an isolated traffic-actuated two-phase traffic signal (Morris & Pak-Poy 1967)

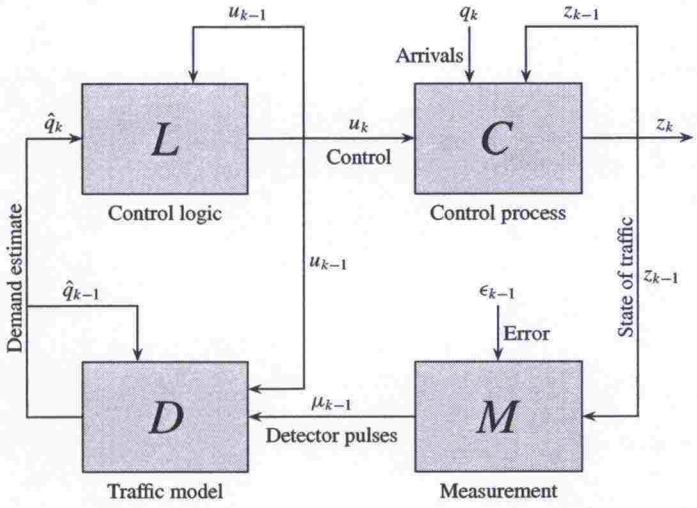


Figure 3.5: Demand-responsive traffic signal control process

Grafton & Newell (1967) used dynamic programming (Bellman 1957) to determine optimal policies for the control of an intersection of two traffic streams using total delay as the objective function to be minimized. They found that the zero-switch queue strategy had to be modified if the initial queues were very large or very small.

According to Newell (1969) at a traffic-actuated signal of two one-way streets the average delay per car is less than for a pretimed signal by about a factor of 3. To minimize delay the signal should (in most cases) switch as soon as the queue vanishes. At an intersection of two two-way streets the relative efficiency of traffic-actuated control is not high compared with one-way streets (Newell & Osuna 1969). Also, the zero-switch queue strategy is very inefficient, and a compromise strategy is required.

Dunne & Potts (1964, 1967) have presented a linear control algorithm based on queue

lengths. The optimum policy is to change the signal when a queue vanishes. This approach may be called a zero-switch queue strategy. If the queue discharge gaps are nearly uniform the zero-switch queue strategy can be implemented as a gap-seeking algorithm with unit extension interval approximately equal to the queue discharge gap.

For oversaturated conditions Gordon (1969) has suggested a control algorithm which keeps the ratio of actual queues to the maximum link storage space on each phase equal. This consideration is important especially in urban conditions. Koshi (1979) has suggested an algorithm that under heavy traffic and oversaturated conditions keeps the degrees of saturation in all phases uniform. Gazis & Potts (1965) suggested a traffic-responsive modification for their fixed-time control of oversaturated intersections.

One of the most advanced implementation of the gap-seeking algorithm is the Swedish LHOVRA method. The name is a Swedish acronym for the system functions truck priority (L—lastbilsprioritering), major roadway priority (H—huvudledsprioritering), accident reduction (O—olycksreduktion), variable yellow change interval (V—variabelt gröngult), extend red clearance (R—rödkörningskontroll), and omit red clearance (A—allrödvändning) (Vägverket 1983, Peterson, Bergh & Steen 1986).

With an extensive set of detectors (Fig. 3.6) the control strategy has functions to:

L: Give priority to trucks, public transport, and platoons

H: Improve traffic quality on the major roadway

O: Reduce the number of vehicles in the option zone at the end of the green phase

V: Shorten the yellow change interval if no vehicles are approaching

R: Extend red clearance inteval in case of apparent threat of incident and for left turners waiting inside the intersection

A: Omit red clearance interval on new demand if there is no conflicting demand (change directly from yellow to green without becoming red)

With an advanced detector configuration and signal group control LHOVRA performs well even compared with optimizing control strategies such as MOVA (Kronborg 1992).

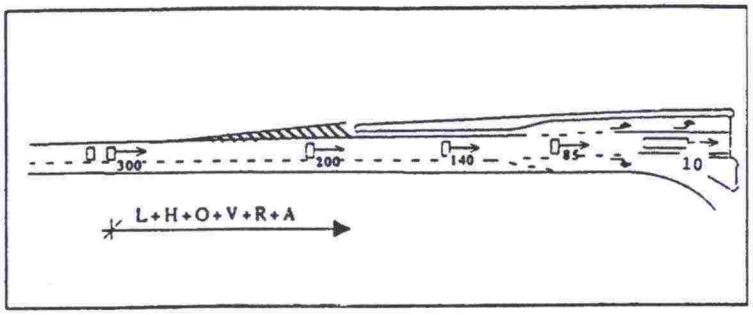


Figure 3.6: Typical LHOVRA detectors in main road (70 km/h) (Kronborg 1992)

LHOVRA has been designed for the signal control of isolated intersections on high speed roads. The most distinctive feature of LHOVRA is that it considers safety as the most important factor in traffic signal control. LHOVRA improves safety and the quality of traffic, especially on the major road. LHOVRA has been a used as a pattern for the current Finnish control strategy described in the next section.

Because the control methods described above do not use effectively all the available information, because variation in queue discharge gaps makes the implementation of theoretically optimal parameters inefficient, and because the optimal traffic-responsive control strategies are different for different traffic conditions, research was early directed toward *optimizing strategies*. See Heydecker (1990) and Bell (1990) for short reviews. The structure of the process is similar to the one presented in Figure 3.5). As compared to the simple demand-responsive methods described above, the traffic models and control logics in optimizing strategies are more advanced.

The algorithm of Miller (1965) calculates at regular intervals delay estimates for changing the signals and for keeping the current state of signals. The signals are left unaltered unless it appears that the least delay will result from an immediate change. A detector placed about 100 m (300 ft) back from the stop line gives arrival information about 8–10 seconds in advance. van Zijverden & Kwakernaak (1969) have described a similar control method based on the minimization of estimated future delays. Bång (1976) extended the objective function to include the cost of both delays and stops. The British MOVA (Vincent & Young 1986, Peirce & Vincent 1989) and Swedish SOS (Kronborg, Davidsson & Edholm 1997) systems are further developed applications of the Miller algorithm.

In the last decades the details is queue, platoon, coordination, filtering, and prediction modeling in adaptive urban traffic control have been further developed (Baras, Dorsey & Levine 1979, Baras, Levine, Dorsey & Lin 1979, Baras, Levine & Lin 1979, Olszewski 1990). Optimization software, such as Transyt (Robertson 1969), has been applied in some traffic-responsive algorithms for urban traffic control (Baras & Levine 1979). Another approach has been the application of dynamic programming (Bellman 1957).

The best known implementation of an optimized coordinated signal control system based on cyclic flow profiles and Transyt-type optimization is SCOOT (Hunt, Robertson, Bretherton & Winton 1981, Hunt, Robertson, Bretherton & Royle 1982*a*, Hunt, Robertson, Bretherton & Royle 1982*b*). SCOOT has also been installed on isolated intersections (Carden & McDonald 1985). Like the British SCOOT, also the Australian SCAT method (Sims 1979, Lowrie 1982) adopts frequent but small changes in control parameters to meet fluctuating traffic demand. Unlike SCOOT, which is a centralized system, SCAT allows phase optimization in local traffic-actuated controllers (Luk 1984).—Rosdolsky (1973) suggested a local control algorithm for global signal coordination. He assumed the results to be valid for light traffic, but did not claim the algorithm to be suitable for implementation in the field.

Demand-responsive control algorithms are implementations of optimization problems. Michalopoulos & Stephanopoulos (1979) has suggested a delay minimizing algorithm with queue-length constraints for critical intersections, which saturate frequently. Dynamic programming with some simplifications has been suggested in many published research reports (Betró, Schoen & Speranza 1987, Gartner 1983, Chen, Cohen, Gartner & Liu 1987). A continuous-time formulation of optimal group-based traffic-responsive control as a non-linear binary mixed integer programming problem has been presented by Heydecker (1990). Also fuzzy control of traffic signals has received attention and some applications in recent years (Pursula & Niittymäki 1996, Niittymäki 1998, Mäenpää 2000, Niittymäki 2002).

For a recent overview of traffic signal control strategies see Smith, Clegg & Yarrow (2001). An overview from a Finnish perspective has been presented by Eloranta (1998).

As this discussion has revealed, numerous optimizing control methods have been im-

plemented, and even more suggested in research papers. Also, the current implementations of gap-seeking algorithms, such as LHOVRA, are very advanced. This presents a challenge for a performance analysis of traffic-responsive signal control. Before a discussion of these methodologies it is appropriate to give a more detailed description of the Finnish traffic signal control practice.

3.2.3 Traffic-responsive control in Finland

In Finland, the most usual controller type is a fully actuated controller, especially at isolated intersections. Green intervals for each signal group are started and extended by the current demand. The actuated controller receives detection every time a vehicle passes a *detector*. When a controller receives information of approaching vehicle, it extends the green phase with a unit extension interval, which is defined separately for each detector. The unit extension is defined in such a way that the approaching vehicle has enough time to get to the next detector. When passing the second detector, a new extension is started.

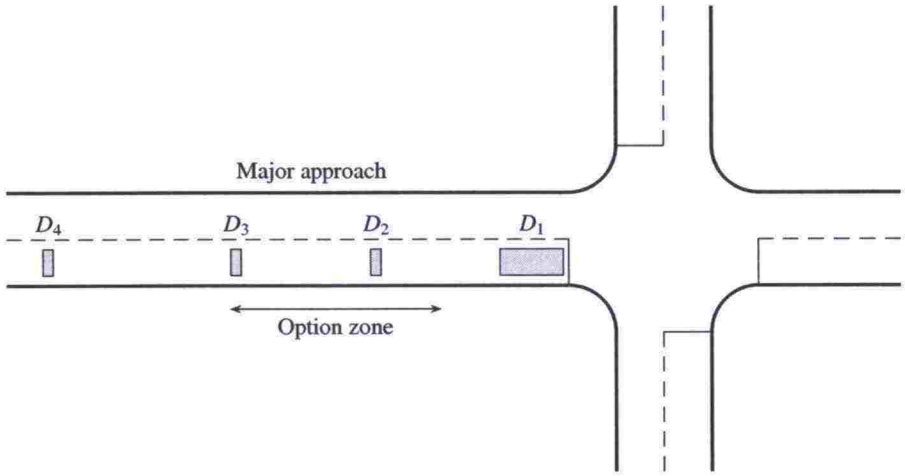


Figure 3.7: Typical detector configuration on the major approach of an isolated intersection

A green interval, once it is started, lasts at least a minimum green time and does not extend beyond a maximum green time, assuming that there is demand for green of conflicting signal groups. Minimum green time is usually related to the safety aspects (for example time needed for pedestrians to cross the street at the same phase). Maximum green time is needed to avoid too long cycle lengths during high demand.

The functions of traffic-responsive control are divided into two groups: minimum functions and quality improving functions (Kehittämiskeskus 1996, Tervala & Appel 1987). Minimum functions improve safety and guarantee an acceptable quality of progression. As far as possible, these functions are implemented at every intersection having traffic-responsive control. The *minimum functions* are:

1. Queue discharge
   (a) The green interval is extended so that the queue accumulated during the red time can discharge during the following green time. Green is extended until the time gap between vehicles indicates that the queue has discharged or until the maximum green is reached.
   (b) Executed by a passage+presence detector (D_1) near the stop line.

(c) This function is a minimum requirement for quality of flow.

2. Option zone clearance

   (a) Option zone is the area of upstream direction, where two drivers may decide to stop or continue, when the green is terminated. Different decisions may cause rear-end collisions, which can be reduced by clearing the option zone from approaching vehicles.

   (b) Accomplished by monitoring the arriving vehicles with passage detectors at the beginning of the option zone (D_3) and in the option zone (D_2). See Figure 3.8.

   (c) Main requirement of traffic safety.

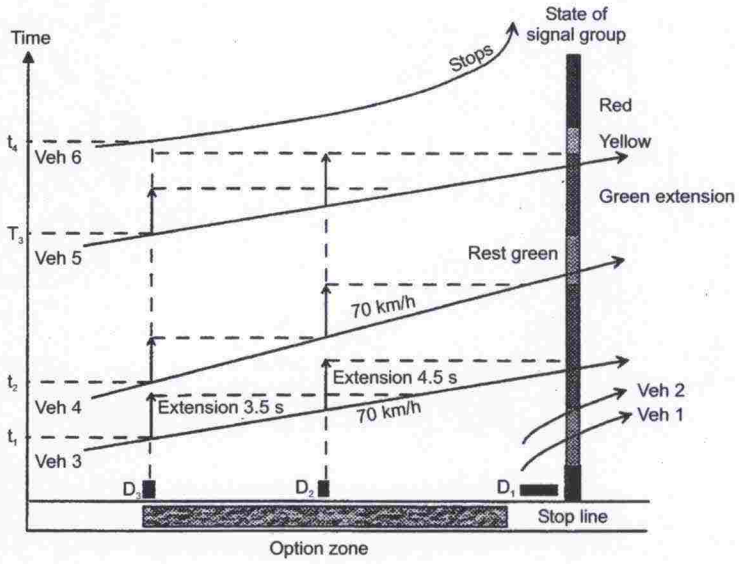


Figure 3.8: Option zone clearance using two passage detectors. Adapted from Kehittämiskeskus (1996).

3. Initial green extension

   (a) A driver, who sees the signal switch to green, may presume that he/she can pass the intersection within this phase. However, short minimum green may terminate before the driver reaches a detector, and the risk of red-run offences increases.

   (b) Executed by the most distant detector (D_4) from the stop line (Figure 3.9).

   (c) Safety requirement on high speed roads (60-70 km/h).

4. Short red-rest interval prevention

   (a) Actuated signal control initiates a red rest mode if no demand (detection) exists neither on the observed signal group nor its conflict groups. To prevent a vehicle from stopping because of a red rest interval, the green phase is started soon enough before the vehicle has to brake.

   (b) Executed by a detector (D_4) far enough from stop line (Figure 3.10). Speed limit and the time required to return back to green define the distance of the detector.

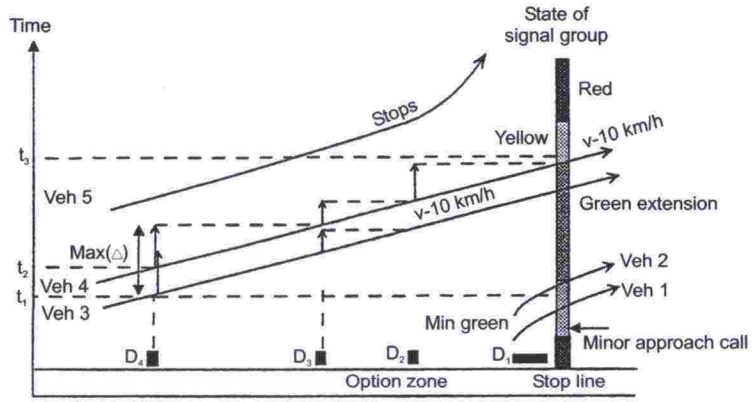


Figure 3.9: Initial green extension. Adapted from Kehittämiskeskus (1996).

(c) Quality-of-flow and safety requirement during low traffic demand.

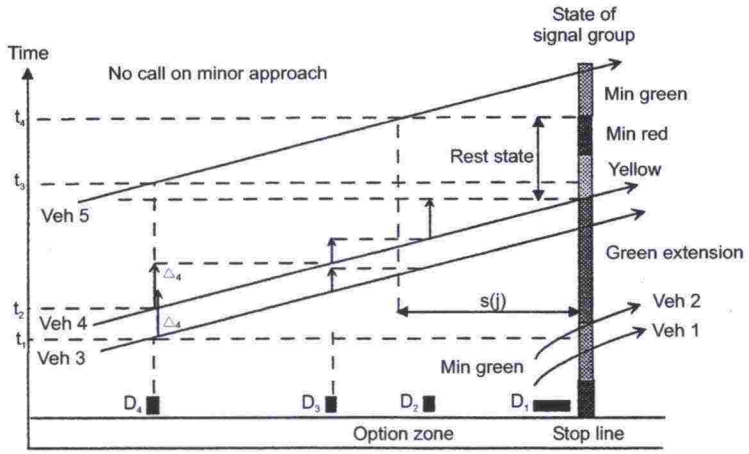


Figure 3.10: Detectors for red rest type control. Adapted from Kehittämiskeskus (1996).

Figure 3.11 displays typical detector locations for the minimum functions.

The *quality-improving functions* are mostly used on high-class arterial roads, like arterial streets and beltways of major cities. In addition to (or even instead of) the average delays, the quality-improving functions try to decrease the number of vehicle stops especially on the major direction. The quality-improving functions are:

1. Green reservation

   (a) When a vehicle passes the most distant detector (D_4), the green for the major flow is reserved (Figure 3.12). The signal switches to green, when the vehicle reaches the option zone (D_3). During the reservation period, the minor roadway green will not be initiated.

   (b) Vehicles behind detector D_4 must not reach breaking distance during a minimum green in the minor roadway.

   (c) Improves quality of flow on the major roadway during low traffic demand.

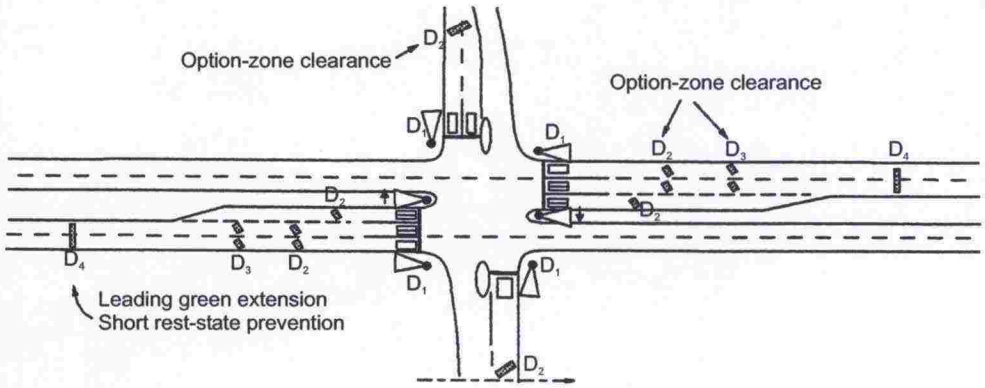


Figure 3.11: Detectors for the minimum functions. Adapted from Kehittämiskeskus (1996).

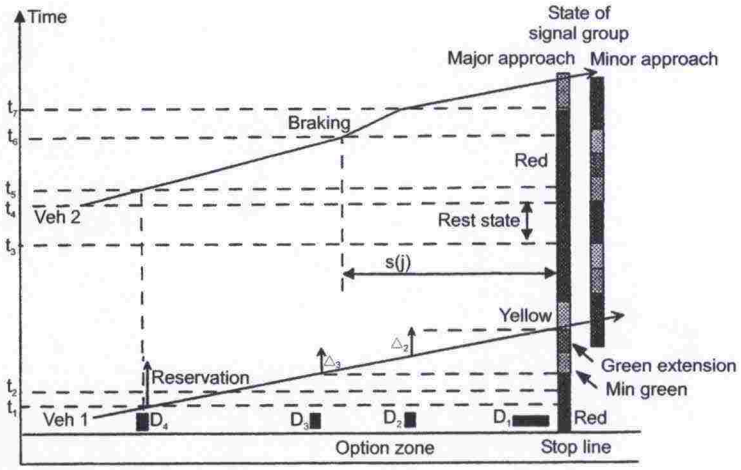


Figure 3.12: Green phase reservation for the major roadway. Adapted from Kehittämiskeskus (1996).

2. Green extension for approaching platoons

   (a) The green phase is extended, if there is three or more vehicles arriving within a six-second time gap (Figure 3.13). If "option zone clearance" is in effect, the extension for platoons is not initialized.

   (b) Executed by the most distant major roadway detectors (D_4).

   (c) Reduces the percent stopped vehicles on the major roadway.

3. Heavy-vehicle and public-transport priority

   (a) Heavy vehicles are detected so far that they do not reach the breaking distance during minimum minor green interval. Depending on the moment of HV-detection, the green interval is extended or started earlier. For public transport, an extra phase can also be added to the cycle.

   (b) Executed by special detectors (D_4) in the major roadway.

   (c) Reduces the stops and delays of heavy vehicles, including buses. Reduces delays for other vehicles on the major roadway by reducing the number of heavy vehicles slowly accelerating from the stop line.

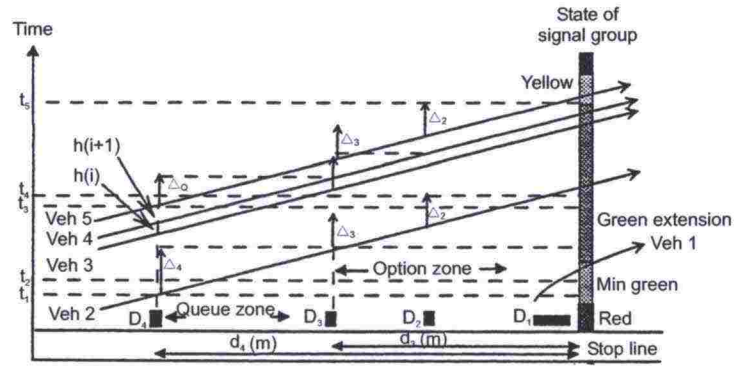


Figure 3.13: Green extension for approaching platoons. Adapted from Kehittämiskeskus (1996).

4. Directional termination of green

   (a) After a specified green time the major roadway signal groups are switched red independently. An extension on one direction does not extend the green on the other direction. Makes safe green phase termination easier, which reduces the number of realized maximum greens and shortens the average cycle lengths.

   (b) Reduces the delays of minor flow by shortening the cycle length and improves safety on major direction.

5. Yellow change interval extension

   (a) If the option zone has no vehicles that would cross the stop line during the yellow change interval, a very short yellow signal is displayed (Figure 3.14).

   (b) Executed by detectors at the beginning and at the end of option zone (D_2 and D_3).

   (c) Reduces the lost time between green phases.

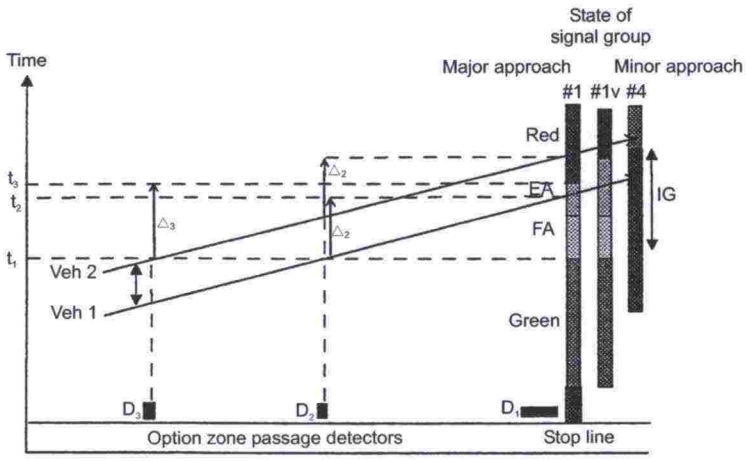


Figure 3.14: Extension of yellow change and red clearance intervals. Adapted from Kehittämiskeskus (1996).

6. Red clearance extension
   (a) Intergreen time is extended, if the option zone detector (D_2) closer to the stop line detects a vehicle during the yellow change interval (Figure 3.14).
   (b) Safety function that reduces the risk of red-run accidents.
7. Different delay, stopping and cost optimization algorithms, not in common use.

Figure 3.15 displays typical detectors for both minimum and quality-improving functions. Other special features of Finnish signal controllers are parameters such as

- *Green initialization.* Options: own (detector) request, fixed request, request from another signal group, request from another signal group in the same phase, initialize if no conflicts, special request.
- *Green initialization offset.* Green initialization is delayed until a specified signal group has switched green. If the leading group does not have a request as the delayed group turns green, no initialization offset is executed. If the initialization offset is *prohibitive*, the green initialization for the leading group is not allowed if the request occurs while the delayed group is already green. If the initialization offset is *permissive*, the leading group is allowed to initialize green, even when the request occurs during the green of the delayed group.
- *Initialization of maximum-green counting.* Options: green start, conflicting request during the green, phase start, conflicting request during the phase.
- *Green termination after detector extensions.* Options: Stay green until the beginning of conflicting green, terminate at conflicting demand, terminate immediately after extensions, extend to maximum green.
- *Termination extension*; i.e., additional extension at the end of green. Options: fixed extension, traffic-actuated extension, fixed + traffic-actuated extension.
- *Rest state.* Options: green, no change, red.

Other parameters include the integreen times, minimum and maximum greens, and detector functions, such as unit extension and gap-reduction parameters.

In Finland traffic-responsive control, with some limitations, is used also in coordinated control (Sane 1986, Luttinen 1994). All functions cannot be applied, because coordinated signals require more or less fixed green starts and similar cycle and green times between intersections. The main principle is to set fixed requests for green in the coordinated (major) direction, while the minor direction gets green only when true demand exists. The initialization of green on the major roadway is restricted to a short time period. Green termination is adjusted by green extensions within a larger time period, but with more limited freedom than in isolated control. In other words, the minor roadway green is started only if demand exists and terminating the major roadway green will not disturb the coordination of major flow. The queue discharge and option zone clearance functions are used. Platoon extensions and heavy-vehicle priorities may also be used, if necessary and possible.

In recent years fuzzy control methods have been developed (Pursula & Niittymäki 1996, Niittymäki 1998, Niittymäki 2002). The research has been mainly focused on the development of control algorithms and their evaluation by simulation with a few field implementations (Mäenpää 2000).

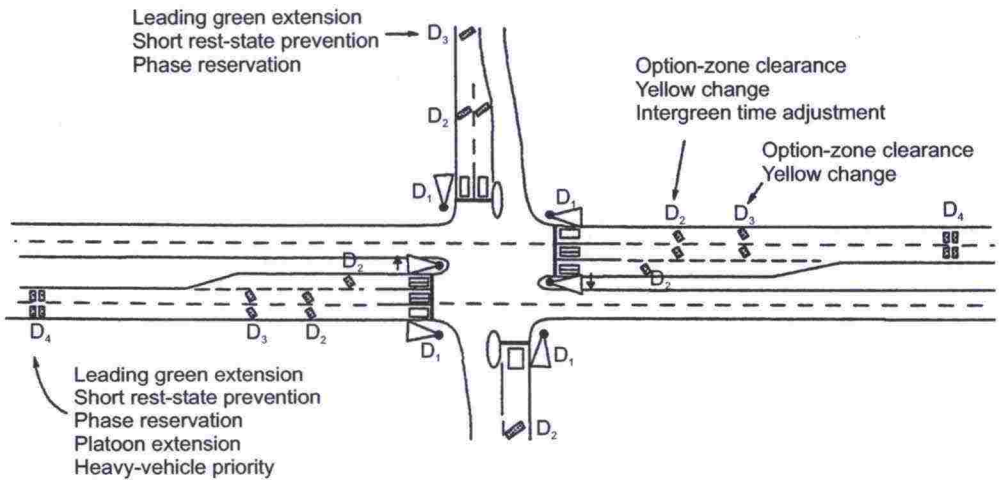


Figure 3.15: Detectors for minimum and quality-improving functions. Adapted from Kehittämiskeskus (1996).

4 DELAY ESTIMATION

4.1 Definitions

The most important performance measure for traffic signal control is delay. It can be defined in various ways, the most common of which are (Fig. 2.5, page 28):

1. Stop delay
2. Time-in-queue delay
3. Control delay

Stop delay is the time that a vehicle spends stopped and waiting for the red signal to turn green. *Time-in-queue delay* starts when a vehicle stops at the end of a queue and ends when the vehicle passes the stop line.—Leutzbach & Köhler (1974) called the three delays *i*) stopped time, *ii*) waiting time, and *iii*) operational delay time.

The delay due to the traffic signal control is called *control delay*. It is the difference between the actual time taken for a vehicle to traverse a road section affected by traffic signals and the time it would have taken to traverse the same road section if there would have been a green signal and no queuing at the intersection (fig. 2.5). The control delay consists of three components:

1. Deceleration delay (W_d)
2. Stop delay (W_s)
3. Acceleration delay (W_a)

Deceleration (acceleration) delay is the time lost due to deceleration (acceleration). A vehicle in a queue may also have an additional delay due to the queue move-up time.

Geometric delay is the time lost due to the intersection geometry. If there is no control delay, geometric delay is the time lost due to deceleration and acceleration required when traversing the intersection, as compared to traversing a comparable straight road section. Geometric delays may be large for turning movements.

Geometric delay depends on the control delay: The deceleration due to signal control or yielding to other traffic streams may compensate the deceleration required by the intersection geometry. If a vehicle has to stop at red signal, no additional deceleration due to intersection geometry is required.—Total delay of a vehicle is the sum of control delay and geometric delay.

The performance of a signalized intersection is often described in terms of level of service (LOS). The LOS classification in HCM has six categories (A-F), which are defined by the average control delay per vehicle. Category A describes a situation with extremely favorable traffic flow progression. Category F describes unacceptable conditions, which usually occur when the traffic demand exceeds capacity ($\rho > 1.0$). The critical delays of each category according to Finnish Traffic Signal Handbook (Kehittämiskeskus 1996) and HCM2000 are presented in table 4.1.

The critical values in HCM2000 are control delays. Finnish guidelines follow the critical values in the 1985 HCM (Transportation Research Board 1985), which uses average stop delay as a service measure. Stop delay is assumed to be 77 % of the control delay (Hurdle 1984, Akçelik 1988).

Table 4.1: Level-of-service criteria according Finnish and American planning standards

Level of service	Delay per vehicle (s)	
	Finnish guidelines	HCM2000
A	≤ 5	≤ 10
B	5–15	10–20
C	15–25	20–35
D	25–40	35-55
E	40-60	55-80
F	> 60	> 80

4.2 Fluid analogy model

4.2.1 Equilibrium conditions

Let us assume that vehicles arrive at an intersection at arrival rate $q(t)$. The cumulative number of arrivals is

$$A(t) = \int_0^t q(u)\,\mathrm{d}u. \tag{4.1}$$

If the arrival rate is constant, the cumulative arrival curve is linear: $A(t) = qt$ (Fig. 4.1). Cumulative demand is the number of arrivals plus the initial queue at the beginning of the observation period (Gazis & Potts 1965):

$$A'(t) = L(0) + A(t). \tag{4.2}$$

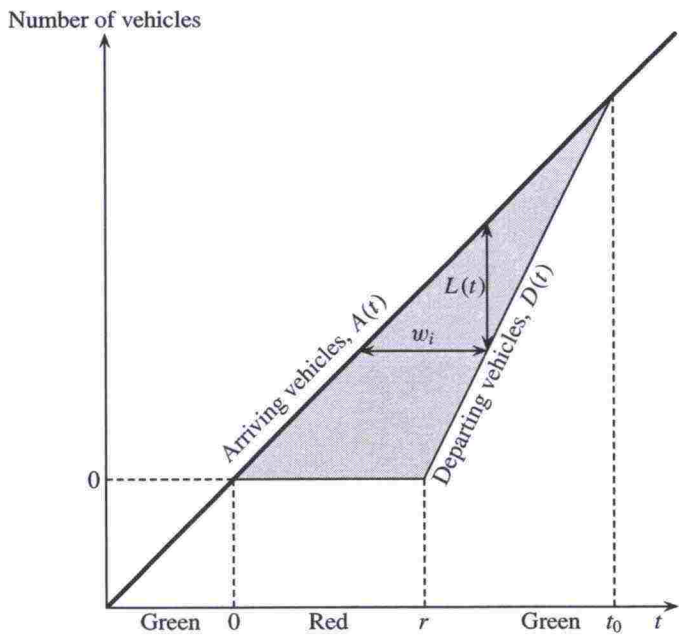


Figure 4.1: Fluid analogy model for traffic signal control

The cumulative number of departures is

$$D(t) = \int_0^t d(u)\,\mathrm{d}u, \tag{4.3}$$

where $d(t)$ is the departure rate. During the red interval of length r no vehicles depart the intersection; that is

$$\begin{aligned} d(t) &= 0, \qquad 0 < t \leq r \\ D(t) &= 0, \qquad 0 < t \leq r. \end{aligned} \tag{4.4}$$

When the effective green starts at $t = r$, the queue starts to discharge at saturation flow rate s:

$$\begin{aligned} d(t) &= s, && r < t \leq t_0 \\ D(t) &= \int_r^t s \, \mathrm{d}u = s(t - r), && r < t \leq t_0, \end{aligned} \tag{4.5}$$

where t_0 is the time when the queue has discharged. After t_0, vehicles discharge without delay:

$$\begin{aligned} d(t) &= q(t), \qquad t_0 < t \leq c \\ D(t) &= A(t), \qquad t_0 < t \leq c, \end{aligned} \tag{4.6}$$

assuming that $t_0 < c$ and $q(t) \leq s$.

The length of the queue (in vehicles) is the initial queue $L(0)$ at $t = 0$ plus the difference between cumulative arrivals and departures. Assuming a constant arrival rate $q \leq sg/c$ the queue length is (Newell 1965)

$$L(t) = L(0) + A(t) - D(t) = \begin{cases} L(0) + qt, & \text{if } 0 < t \leq r \\ L(0) + qt - s(t - r), & \text{if } r < t \leq t_0 \\ 0, & \text{if } t_0 < t \leq c. \end{cases} \tag{4.7}$$

This is a simplification of reality. The vehicles are assumed to form a vertical stack on the stop line. Real queues are always longer than predicted by this model, because vehicles reach the end of queue before they reach the stop line (Hurdle 1984).

The queue has discharged at time t_0, when the number of arrivals is equal to the number of departures (Newell 1965)

$$L(t_0) = L(0) + A(t_0) - D(t_0) = L(0) + qt_0 - s(t_0 - r) = 0; \tag{4.8}$$

that is,

$$t_0 = \frac{L(0) + sr}{s - q}. \tag{4.9}$$

The number of stops per cycle is

$$N_S(c) = L(0) + A(t_0) = L(0) + qt_0. \tag{4.10}$$

(cf. Cronjé 1983).

The capacity of the approach is the saturation flow rate multiplied by the effective green-time fraction:

$$C = s\frac{g}{c}. \tag{4.11}$$

The system is in *equilibrium* if $L(0) = 0$ and $t_0 \leq c$, that is $q \leq C$, in which case the queue discharges before the effective red interval begins. All cycles are similar, so that $L(t) \equiv L(t \bmod c)$.

During a differential time interval $(t, t + \mathrm{d}t]$ the total delay is $L(t)\,\mathrm{d}t$. The total uniform delay during the cycle length c is equal to the area of the triangle in figure 4.1

(Newell 1965)

$$W_u(c) = \int_0^c L(t)\,dt = \int_0^{t_0} qt\,dt - \int_r^{t_0} s(t-r)\,dt$$
$$= \frac{qsr^2}{2(s-q)} = \frac{r^2}{2\left(\frac{1}{q} - \frac{1}{s}\right)}, \qquad q \le \frac{sg}{c}. \tag{4.12}$$

The average uniform delay per vehicle can be obtained by dividing the total uniform delay by the number of arriving (and departing) vehicles during the cycle (Newell 1965):

$$w_u = \frac{\int_0^c L(t)\,dt}{\int_0^c q(t)\,dt} = \frac{W_u(c)}{qc} = \frac{r^2}{2c\left(1 - \frac{q}{s}\right)} = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \frac{g}{c}\rho\right)}, \qquad \rho \le 1, \tag{4.13}$$

where

$$\rho = \frac{qc}{sg} = \frac{q}{C} \tag{4.14}$$

is the demand-to-capacity ratio; i.e., degree of saturation. Accordingly, the average delay is a function of the cycle length (c), the degree of saturation (ρ), and the proportion (g/c) of effective green in the cycle.

4.2.2 Oversaturated conditions

If arrival rate exceeds capacity, the intersection is oversaturated. The number of arrivals during cycle $i \in \mathbb{N}$ is $A_i = A[ic] - A[(i-1)c]$. For a constant arrival rate $A_i = qc$. The departure rate is s during the entire green interval, so that $D(t+c) - D(t) = D(c) = sg$.

If the queue length at the end of cycle i is L_i, the queue at the end of the next cycle is $L_{i+1} = L_i + A_{i+1} - D(c) = L_i + A_{i+1} - sg$ (Akçelik 1980). Because in oversaturated conditions $A_{i+1} > D(c)$, the system is in a *nonequilibrium* state, and the queue length increases cycle by cycle, $L(t+c) > L(t)$.

Figure 4.2 displays a case where demand flow rate instantaneously increases above the capacity at the beginning of a cycle. The capacity curve $C(t)$ is not the saw-toothed departure curve $D(t)$, but a straight line with slope C, so that the area between $A(t)$ and $C(t)$ curves is the overflow delay (Gazis & Potts 1965).

The total delay has two components: uniform delay (gray triangles between capacity and departure curves) and overflow delay (light gray area between arrival and capacity curves). For the estimation of overflow delay only the initial queue $L(0)$, arrival flow rate $q(t)$, and capacity C need to be known. A special strength of this approach is that the same approach can be applied to the estimation of overflow delay at different kinds of facilities, such as freeways (May & Keller 1967) and unsignalized intersections (Kimber, Marlow & Hollis 1977).

The *average uniform delay* is half of the red interval (Hurdle 1984), as obtained by substituting $\rho = 1$ into equation (4.13):

$$w_u = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \frac{g}{c}\right)} = \frac{r}{2}. \tag{4.15}$$

In a general form the equation can be expressed as (Fambro & Rouphail 1997)

$$w_u = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \min\{\rho, 1\}\frac{g}{c}\right)}. \tag{4.16}$$

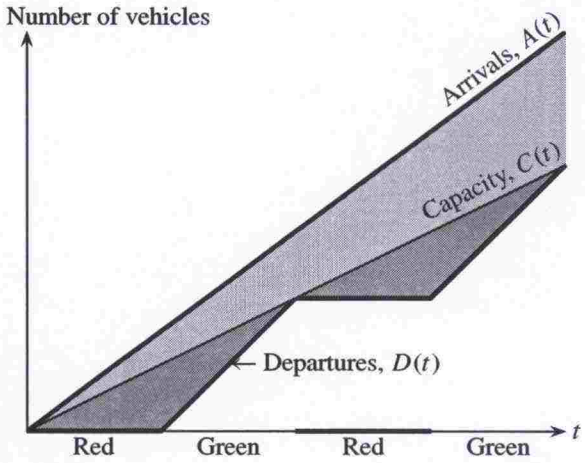


Figure 4.2: Fluid analogy model for overflow delay

The overflow queue is

$$L_o(t) = L(0) + A(t) - C(t), \tag{4.17}$$

assuming that the overflow condition continues during the whole time interval (cf. Catling 1977):

$$\forall u \in (0, t]: L(0) + A(u) - C(u) > 0. \tag{4.18}$$

If the overflow period starts at $t = 0$, there is no initial queue, and $L_o(t) = A(t) - C(t)$. For a constant arrival rate and capacity the overflow queue is $L_o(t) = t(q - C) = Ct(\rho - 1)$. If the overflow queue at time instant t_1 is $L_o(t_1)$, the overflow queue at $t_2 > t_1$ is

$$L_o(t_2) = L_o(t_1) + A(t_2) - A(t_1) - C(t_2) + C(t_1) = L_o(t_1) + C(\rho - 1)(t_2 - t_1), \tag{4.19}$$

assuming constant arrival rate and capacity. Under condition (4.18) the equation holds also for $\rho < 1$.

The overflow delay (Figure 4.3) during interval $(0, t]$ is the area between the cumulative arrival curve and the capacity curve (Gazis & Potts 1965, May & Keller 1967):

$$W_o(t) = \int_0^t L_o(u)\,\mathrm{d}u = \frac{t^2}{2}(q - C) = \frac{Ct^2}{2}(\rho - 1). \tag{4.20}$$

This delay includes the overflow delay accumulated until t, that is the overflow delay 0PQ of vehicles $C(t)$ and the overflow delay PQR of vehicles $A(t) - C(t)$ accumulated up to time t. Because the delay is based on the measurement of queue lengths, this approach is called the *queue-sampling method.* Delay estimation based on the measurement of individual vehicle delays is called the *path-trace method* (Rouphail & Akçelik 1992).

The average overflow delay per vehicle (Fig. 4.4) based on queue sampling is obtained by dividing the above equation by Ct (Neuburger 1971, Hurdle 1984):

$$w_o(t) = \frac{W_o(t)}{Ct} = \frac{t}{2}(\rho - 1). \tag{4.21}$$

The method overestimates the average overflow delay, because it includes also the delay PQR.

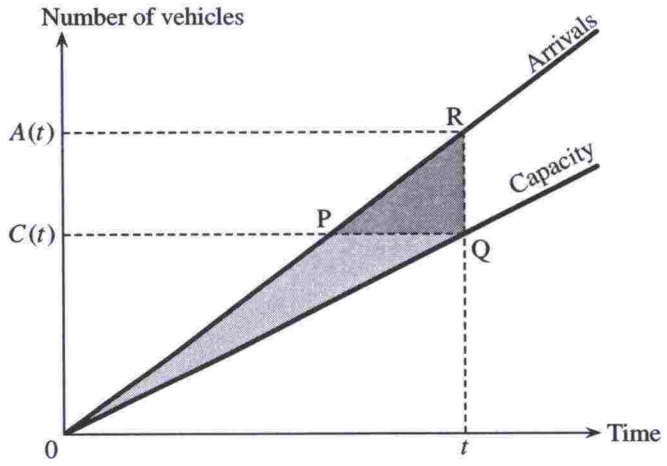


Figure 4.3: Overflow delay based on queue sampling

The average delay is directly proportional to t. As the length of the oversaturation period increases, the curve becomes steeper, and in the limit becomes vertical, as time approaches infinity. The average overflow delay during time interval $(t_1, t_2]$ is

$$w_o(t_1, t_2) = \frac{\int_{t_1}^{t_2} L_o(u)\,\mathrm{d}u}{C(t_2 - t_1)} \tag{4.22}$$

$$= \frac{(t_2^2 - t_1^2)(q - C)}{2C(t_2 - t_1)} \tag{4.23}$$

$$= \frac{t_1 + t_2}{2}(\rho - 1)\,. \tag{4.24}$$

The overflow queue length at t_1 is $L_o(t_1) = Ct_1(\rho - 1)$. If the initial overflow queue length is $L_o(0)$ and condition (4.18) holds, the average overflow delay during $(0, t]$ is obtained from the equation above as

$$w_o(t) = \frac{L_o(0)}{C} + \frac{t}{2}(\rho - 1). \tag{4.25}$$

Rouphail & Akçelik (1992) calculate the average overflow delay per arriving vehicle, which gives

$$w_o'(t) = \frac{W_o(t)}{qt} = \frac{t(\rho - 1)}{2\rho}. \tag{4.26}$$

This method underestimates the average delay, because only part of the delay is attributed to the $Ct(\rho - 1)$ vehicles.

At oversaturated conditions the queues increase in length cycle by cycle, and the system does not reach equilibrium. In the extreme, after infinite time, there will be an infinite queue. The analysis of equilibrium conditions is apparently implausible, when the system is oversaturated. There will be available neither infinite time for queue accumulation nor infinite space for the queues. Severe oversaturation has also an impact on traffic demand, as some traffic redistributes into the surrounding network an avoids the oversaturated intersection (Kimber et al. 1977). Consequently, oversaturation is always a peaked phenomenon lasting only for a limited length of time (Neuburger 1971, Yagar 1977).

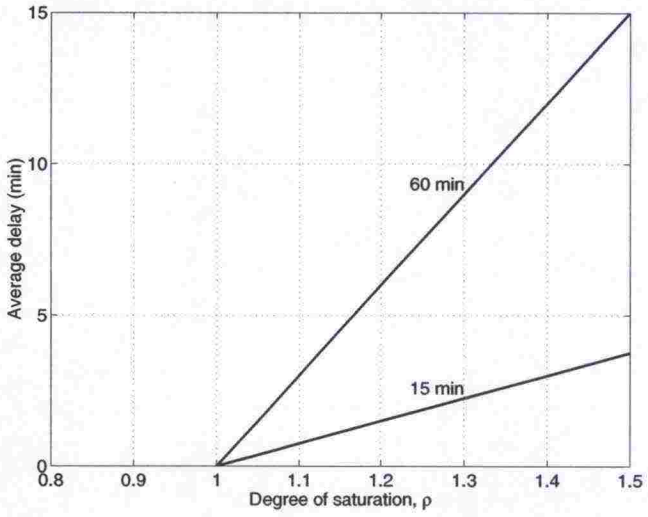


Figure 4.4: Average overflow delay for oversaturation periods of length 15 and 60 minutes

The discussion above provides some tools for the analysis of an oversaturation during a peak period. In order to estimate the delay during a peak period we need to know the maximum flow rate and the shape of the demand pattern. May & Keller (1967) suggested triangular or trapezoid shaped demand patterns (Fig. 4.5). Kimber & Hollis (1978) defined a low-definition (rectangular) and a high-definition (piecewise constant) pattern.

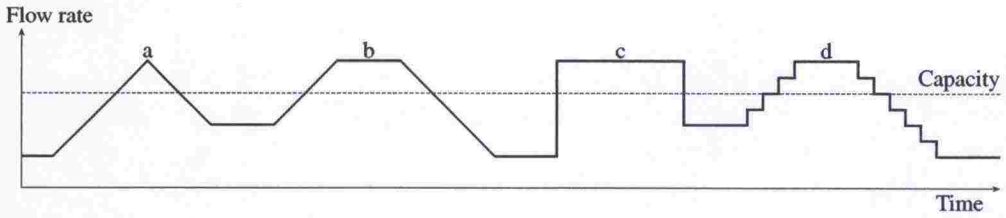


Figure 4.5: Triangular (a), trapezoidal (b), rectangular (c), and piecewise constant (d) demand patterns

Let us assume a rectangular demand pattern with peak flow $q_p > C$ lasting for time t_p, after which the flow rate decreases to q. The time $(0, t_p]$ is the *peak period.* There is no initial queue, $L(0) = 0$, in the system. During the time interval $(0, t_p)$ the overflow queue increases at rate $C(\rho_p - 1)$ and reaches its maximum $L_o(t_p) = Ct_p(\rho_p - 1)$, where $\rho_p = q_p/C > 1$ is the degree of saturation during the peak period (Fig. 4.6). After t_p the queue length decays at rate $C(\rho - 1)$, where $\rho < 1$ is the degree of saturation after the peak period (Kimber & Hollis 1978). The overflow queue is

$$L_o(t) = \begin{cases} Ct(\rho_p - 1), & \text{if } 0 < t \leq t_p \\ L_o(t_p) - C(t - t_p)(1 - \rho), & \text{if } t_p < t \leq t_o \end{cases} \tag{4.27}$$

The *overflow period* ends at t_o, when the overflow queue vanishes:

$$L(t_o) = L_o(t_p) - C(t_o - t_p)(1 - \rho) = 0. \tag{4.28}$$

The length of the overflow period is determined by the length of the peak period and the degree of saturation during and after the peak period (see Kimber & Hollis 1978):

$$t_o = t_p + \frac{L_o(t_p)}{C(1-\rho)} = \frac{(\rho_p - \rho)t_p}{1-\rho}. \tag{4.29}$$

This result can also be obtained more directly by equating the the number of overflow vehicles to the capacity during time interval $(0, t_o]$ (Rouphail & Akçelik 1992):

$$q_p t_p + q(t_o - t_p) = Ct_o. \tag{4.30}$$

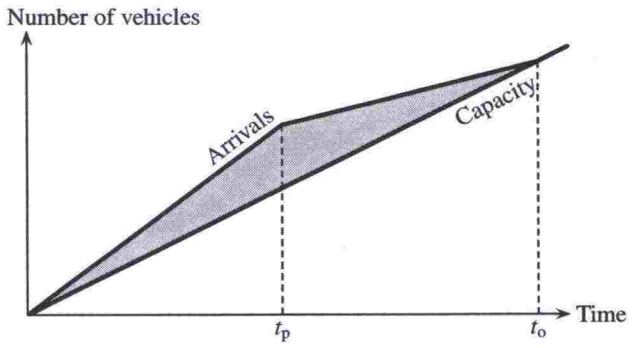


Figure 4.6: Overflow delay due to a rectangular peak flow pattern

Total overflow delay is the area of the triangle in Figure 4.6 (cf. Kimber & Hollis 1978):

$$W_o = \int_0^{t_o} L_o(t)\,\mathrm{d}t = \frac{Ct_p^2(\rho_p - 1)(\rho_p - \rho)}{2(1-\rho)}. \tag{4.31}$$

Since the queue increases and decreases in a linear fashion the average queue length during time interval $(0, t_o]$ is half the maximum queue length of $L_o(t_p)$ (Neuburger 1971), and the delay is

$$W_o = \frac{t_o}{2} L_o(t_p). \tag{4.32}$$

Because the delay includes the area of the total triangle, the results for both queue-sampling and path-trace methods agree, and the bias in the average overflow delay disappears.

The average overflow delay per vehicle is obtained by dividing the total overflow delay with the number Ct_o of vehicles experiencing overflow delay:

$$w_o = \frac{W_o}{Ct_o} = \frac{L_o(t_p)}{2C} = \frac{t_p}{2}(\rho_p - 1). \tag{4.33}$$

The same result can also be derived using the Little's formula, which states that the average queue length is equal to the product of average waiting time and the arrival rate (Little 1961). Because during the interval $(0, t_o]$ the average queue length is $L_o(t_p)/2$, and the average arrival rate to the overflow queue is equal to the departure rate (capacity), we obtain

$$w_o = \frac{L_o(t_p)}{2C} = \frac{t_p}{2}(\rho_p - 1). \tag{4.34}$$

As equation (4.21) shows, this delay is equal to the average queue sampling delay during time interval $(0, t_p]$.

The average overflow delay does not depend on the degree of saturation after the peak. Because the average overflow queue length $L_o(t_p)/2$ does not change with ρ, the delay per vehicle does not increase, although ρ increases. The time period of higher average delays, however, becomes longer.

The equation above gives the average overflow delay per vehicle *during the overflow period* $(0, t_o]$. Kimber & Hollis (1978) have attributed the excess delay to the vehicles $A(t_p) = C\rho_p t_p$ of the peak itself:

$$w_o = \frac{W_o}{C\rho_p t_p} = \frac{t_p(\rho_p - 1)(\rho_p - \rho)}{2\rho_p(1 - \rho)}. \tag{4.35}$$

One should be careful, if this average delay is used to estimate the total overflow delay.

The total average delay during the overflow interval $(0, t_o]$ is the sum of average uniform and overflow delays:

$$w = w_u + w_o = \frac{1}{2}[r + t_p(\rho_p - 1)]. \tag{4.36}$$

Figure 4.7 displays the average deterministic delay per vehicle for a 15-minute peak period when cycle length is 100 s and green interval 50 s. For $\rho \leq 1$ the figure describes steady state conditions. For $\rho > 1$ the observation period is equal to the overflow period $(0, t_o]$. It is assumed that the accumulation rate of the uniform delay is constant. The results are theoretically accurate only when the overflow period is a multiple of the cycle length ($i \in \mathbb{N}$: $t_o = ic$).

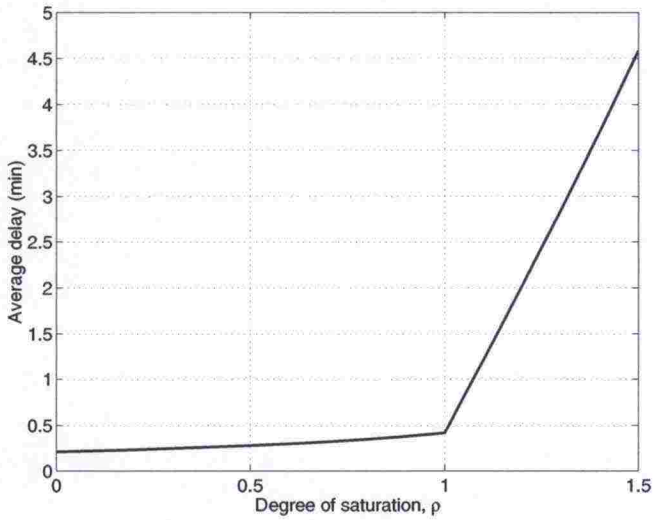


Figure 4.7: Average deterministic delay per vehicle for a 15-minute peak period when cycle length is 100 s and green interval 50 s

If the time period observed $(0, t]$ includes both the entire overflow period and noncongested time periods, the average delay should be calculated as

$$w = \frac{W_u + W_o}{A(t)}. \tag{4.37}$$

If the time period includes only part of the overflow period a conscious decision should be made as to which vehicles the excess delay is attributed. The average delay based

on queue sampling is

$$w = \frac{\int_0^t L_u(y) + L_o(y)\,dy}{A(t)}, \tag{4.38}$$

where $L_u(t)$ is the steady state length of the queue (dark triangles in figure 4.2). The calculations should consider the initial queue at the start of the analysis period. Rouphail & Akçelik (1992) have discussed the delay estimates based on the path-trace method.

4.3 Effect of random arrivals

The fluid analogy model does not consider the random nature of the arrival process. Even under stationary conditions, when the average arrival rate does not change, the number of arrivals during the effective red and green intervals changes from cycle to cycle. At low degrees of saturation this random effect does not play a large role. As the degree of saturation approaches unity, the demand on some cycles exceeds the cycle capacity (Haight 1959). This *random overflow* causes excess delay, which the fluid analogy model does not consider (Hurdle 1984).

Webster (1958) estimated the random effect referring to a M/D/1 queuing system, where the arrival process is Poisson and the customers are served by one server having constant service times. He used the Pollaczek-Khintchine equation as presented in the classic paper of Kendall (1951). According to this equation the expected waiting time of customers in a M/G/1 queuing system with Poisson arrivals (interarrival times follow the negative exponential distribution) and one server with service times following some general distribution is

$$w_G = \frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right), \tag{4.39}$$

where τ is the mean and σ^2 is the variance of service times, $\rho = \lambda\tau$ is the server utilization factor, and λ is the average arrival rate.

In an M/D/1 queuing system the service times are constant ($\sigma^2 = 0$), and the expected waiting time is

$$w_D = \frac{\rho\bar{\tau}}{2(1-\rho)} = \frac{\rho^2}{2\lambda(1-\rho)}. \tag{4.40}$$

In traffic signals the average service time is $\tau = c/sg$, and the utilization factor is $\rho = qc/sg$, which is equal to the demand-to-capacity ratio (q/C). In traffic engineering notation the equation can be expressed as

$$w_D = \frac{\rho^2}{2q(1-\rho)} = \frac{\rho}{2C(1-\rho)}. \tag{4.41}$$

The expected waiting time approaches zero at low degrees of saturation and increases to infinity as the degree of saturation approaches unity. The equation has the form (Li, Rouphail & Akçelik 1994, Fambro & Rouphail 1997)

$$w_D = \frac{\alpha\rho}{C(1-\rho)}, \tag{4.42}$$

where α is an adjustment factor for the effect of randomness. For a uniform arrival process $\alpha = 0$, and for Poisson arrivals $\alpha = 0.5$. Newell (1965, 1989, 1990) has suggested $\alpha = (I_A + I_D)/2$, which is the mean of variance to mean ratios for the arrival and departure processes. In an M/D/1 process $I_A = 1$ and $I_D = 0$, and in the case of uniform arrivals $I_A = I_D = 0$, which produce results consistent with the discussion above.

Webster (Webster 1958, Webster & Cobbe 1966) estimated the average delay in traffic signals at equilibrium conditions as the sum of uniform and random delays (Fig. 4.8):

$$w = w_u + w_D = \frac{c\left(1-\frac{g}{c}\right)^2}{2\left(1-\frac{q}{s}\right)} + \frac{\rho^2}{2q(1-\rho)}. \tag{4.43}$$

He observed, however, that this sum slightly overestimated the control delay. This bias was corrected by a third term, estimated by simulation, leading to equation

$$w = \frac{c\left(1-\frac{g}{c}\right)^2}{2\left(1-\frac{q}{s}\right)} + \frac{\rho^2}{2q(1-\rho)} - 0.65\left(\frac{c}{q^2}\right)^{\frac{1}{3}}\rho^{2+5g/c}. \tag{4.44}$$

The correction term was generally in the range of 5 to 15 percent of the sum $w_u + w_D$. For practical purposes the control delay was approximated as

$$w = 0.9\left[\frac{c\left(1-\frac{g}{c}\right)^2}{2\left(1-\frac{q}{s}\right)} + \frac{\rho^2}{2q(1-\rho)}\right]. \tag{4.45}$$

To make the calculations even simpler Webster (1958) presented the equation as

$$w = cA + \frac{B}{q} - C, \tag{4.46}$$

where the values for A, B, and C were presented in tables.

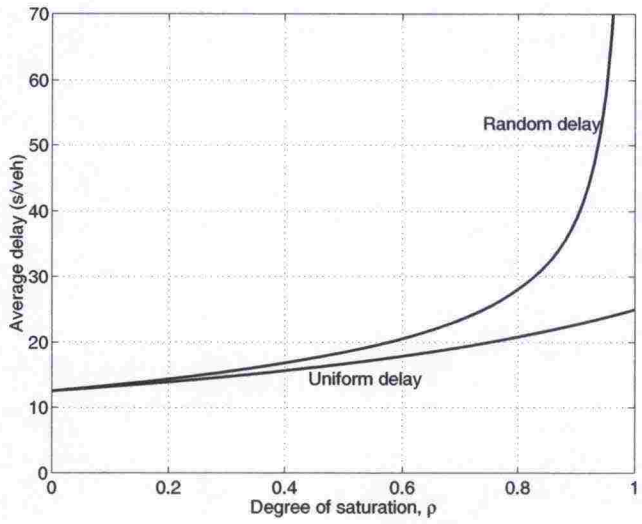


Figure 4.8: Uniform and random delay components for steady state signal control with cycle length 100 s, green interval 50 s, and capacity 1000 veh/h

Webster's formula (4.44) does not have a solid theoretical foundation. It is a sum of average delays from two different processes, namely fluid-analogy model and M/D/1 queuing process, adjusted with a correction term. Many researchers have tried to derive the delay theoretically.

Winsten (Beckmann, McGuire & Winsten 1956) derived the average delay per vehicle in traffic signals with binomial arrivals as

$$w = \frac{c-g}{c(1-p)}\left(\frac{\mathbb{E}[L(0)]}{p} + \frac{c-g+1}{2}\right), \tag{4.47}$$

where $p = q/s$ is the probability of an arrival during a short time interval $\delta = 1/s$, and $\mathbb{E}[L(0)]$ is the expected overflow queue length at the start of a red phase. The cycle length is $c\delta$ and the length of the effective green phase is $g\delta$.

Newell (1960) used the Winsten equation (4.47) and estimated the expected overflow queue at nearly critical arrival rate as

$$\mathbb{E}[L(0)] = \frac{g(c-g)}{2c(g-pc)} \qquad \text{for } p \to \frac{g}{c}. \tag{4.48}$$

For Poisson arrivals he estimated

$$\mathbb{E}[L(0)] = \frac{g}{2(g-pc)} = \frac{g}{2(g-\frac{cq}{s})}, \tag{4.49}$$

so that the average delay estimate for a nearly critical arrival rate was

$$w = \frac{(c-g)^2}{2c(1-p)} + \frac{c}{2(g-pc)} = \frac{(c-g)^2}{2c(1-\frac{q}{s})} + \frac{1}{2\left(\frac{g}{c}-\frac{q}{s}\right)} \tag{4.50}$$

Five years later Newell (1965) approximated the average delay as

$$w = \frac{(c-g)^2}{2c(1-\frac{q}{s})} + \frac{IH(\mu)}{2s\left(\frac{g}{c}-\frac{q}{s}\right)}, \tag{4.51}$$

where $I = I_A + I_D$, as explained above,

$$H(\mu) = \frac{2\mu^2}{\pi}\int_0^{\pi/2} \tan^2\theta \left[-1 + \exp\left(\frac{\mu^2}{2\cos^2\theta}\right)\right]^{-1} \mathrm{d}\theta, \tag{4.52}$$

and

$$\mu = \frac{sg-qc}{\sqrt{Isg}}. \tag{4.53}$$

Darroch (1964) extended Winsten's approach by assuming that vehicles arrive in batches of random size, and the batches are separated by time intervals following the negative exponential distribution. Departures were modeled as a discrete-time process. Darroch derived inequalities for the expected queue length and delay.

McNeil (1968) used a similar compound Poisson arrival model as Darroch (1964). The discharge headways were uniform, and the average delay was estimated as

$$w = \frac{c-g}{2(1-\rho)}\left(c-g+\frac{2\,\mathbb{E}[L(0)]}{q} + \frac{1+\frac{I_A}{1-\rho}}{s}\right). \tag{4.54}$$

Miller (1963) used the results of Winsten to obtain a model for average delay with any variance-to-mean ratio in the arrival pattern:

$$w = \frac{1-\frac{g}{c}}{2(s-q)}\left(\frac{sI_A(2\rho-1)}{q(1-\rho)} + s(c-g) + I_A - 1 + \frac{q}{s}\right). \tag{4.55}$$

The number of arrivals at disjoint time intervals were assumed to be independent.

As this short overview shows, even for pretimed traffic signals many different formulas have been proposed to estimate the average delay. No single method can be assumed to be "correct", not even theoretically. More extensive overviews and comparisons of these and other methods and their modifications have been presented by Rouphail, Tarko & Li (1997) and in the references below.

Allsop (1972) observed that the behavior of real traffic does not correspond closely enough to any of the proposed arrival process models to warrant the very refined solutions that have been obtained. In the comparisons presented by Hutchinson (1972) the differences between the methods by Webster (4.44), Newell (4.51) and Miller (4.55) were shown to be within the limits of accuracy in the measurement of delay and model parameters (cf. Teply 1989). The use of an expression most convenient was suggested. For arrivals with high variability ($I_A > 1$) the Webster model (4.44), however, underestimated the delay. According to Ohno (1978) and Cronjé (1983) Newell's model (4.51) is the most aacurate of the models compared by Hutchinson (1972).

4.4 Delays at traffic-responsive signals

The earliest rigorous analysis of traffic-actuated traffic signals was presented by Garwood (1940). He considered Poisson arrivals to traffic signals with predetermined unit extension time and maximum green period. Darroch, Newell & Morris (1964) also assumed Poisson arrivals, departure headways had any specified distribution, and there was a random lost time for between the signal phases. Minimum and maximum green intervals were not included in the model. They observed that the model was very sensitive to the unrealistic features of the Poisson process. Dunne (1967) and Potts (1967) used a discrete time model with binomial arrivals and zero-switch queue strategy to analyze a two-phase traffic-actuated traffic signal. Newell (1969) and Newell (1969) used fluid and diffusion queuing approximations to evaluate the performance of traffic-actuated traffic signals at nearly congested conditions. Lehoczky (1972) studied zero-switch queue strategy with an arrival process modeled as a Markov chain. For an analysis of semi-actuated traffic signal control see e.g. Haight (1959) and Little (1971).

The theoretical formulations have been complicated and based on significant simplifications of the control process. In particular, the models based on gap-seeking algorithms are very sensitive to the properties of the arrival process (see Luttinen 1996). As an alternative to the theoretical approach a modification of pretimed delay formulas has been suggested.

Courage & Papapanou (1977) estimated the delay in traffic-responsive signals by modifying the Webster equation (4.44). The degree of saturation for the random delay w_D, equation (4.41), was calculated using maximum cycle length and maximum green time:

$$w = w_U + w'_D = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \frac{q}{s}\right)} + \frac{x^2}{2q(1-\rho^*)}, \tag{4.56}$$

where

$$\rho^* = \frac{q c_{max}}{s g_{max}}, \tag{4.57}$$

and c_{max} is the maximum cycle length and g_{max} is the maximum effective green interval for the observed lane group. They ignored the correction term. Michalopoulos,

Papapanou & Binseel (1978) further developed the model as

$$w = w_U + w'_D = b_1 w_U + b_2 w'_D + b_3 \frac{w'_D}{w_U^2}. \tag{4.58}$$

Lin (1983) modified the green proportion g/c and the degree of saturation ρ in the Webster model by new coefficients (b_1, b_2):

$$w = 0.9 \left(\frac{c\left(1 - b_1 \frac{g}{c}\right)^2}{2\left(1 - b_1 b_2 \frac{q}{s}\right)} + \frac{(b_2 \rho)^2}{2q(1 - b_2 \rho)} \right). \tag{4.59}$$

Traffic responsive signals include a large variety of systems. These cannot be easily characterized by one or a few parameters, and the differences between them may well be larger than the systematic difference between pretimed and traffic-responsive control (Swedish National Road Administration 1995).—An extensive literature review is provided by Rouphail, Tarko & Li (1997) as well as in the Appendix A of NCHRP project 3-48 final report (The University of Florida Transportation Research Center 1996).

4.5 Coordinate transformation method

As discussed above, oversaturation is always a peaked phenomenon lasting only for a limited length of time. For a peak period of limited length the average delay does not increase to infinity as the steady-state queuing model indicates. Consequently, the steady-state approach breaks down at high degrees of saturation (Rouphail, Tarko & Li 1997).

At degrees of saturation considerably above unity the queues are very large and the effects of random variation can be ignored (Yagar 1977). The average overflow delay approaches the deterministic overflow delay curve w_o presented above. On the other hand, for degrees of saturation considerably below unity the probability of overflow due to random fluctuations is very low, and the steady-state results appear to be plausible. When the degree of saturation is near unity the average delay is above the fluid analogy model but below the random model (see Hurdle 1984). According to Taale & van Zuylen (2001) the Webster formula (4.44) is valid only up to a degree of saturation of 0.9.

The stochastic nature of a peak period has been studied analytically by de Smit (1971) and Newell (1982). Chodur & Tracz (1984) used simulation. Haight (1963) used a transition matrix (Markov chain) for queue length probabilities. Brilon & Wu (1990) and Wu (1992) have developed the Markov-chain method to estimate delays at pretimed traffic signals under time-dependent conditions. The most commonly used approach is, however, based on a combination of a stationary random model and deterministic oversaturation model.

The average delay curve of the steady-state model approaches asymptotically the vertical line $\rho = 1$ (Fig. 4.8). If it is assumed that the average time-dependent random delay curve should approach the deterministic overflow curve, the new delay curve can be obtained by a coordinate transformation suggested by Kimber et al. (1977) and Catling (1977). Originally the method was applied in the Transyt program (Robertson 1969) by P. D. Whiting.

Let us assume that the degree of saturation is $\rho > 1$. The horizontal difference between the time-dependent delay curve and the deterministic overflow delay curve (w_o) is

$\rho_p - \rho$. The average deterministic overflow delay is

$$w_o(t) = \frac{t}{2}(\rho_p - 1). \tag{4.60}$$

This gives

$$\rho_p = \frac{2w_o(t)}{t} + 1. \tag{4.61}$$

Where the analysis period is either $t \leq t_p$ or $t = t_o$. If $t < t_o$ there is an initial overflow queue at the beginning of the next analysis period.

If the time-dependent delay curve approaches asymptotically the w_o curve in the same way as the random delay curve w_D approaches asymptotically the vertical line $\rho = 1$ (Fig. 4.9), then we have

$$1 - \rho_n = \rho_p - \rho. \tag{4.62}$$

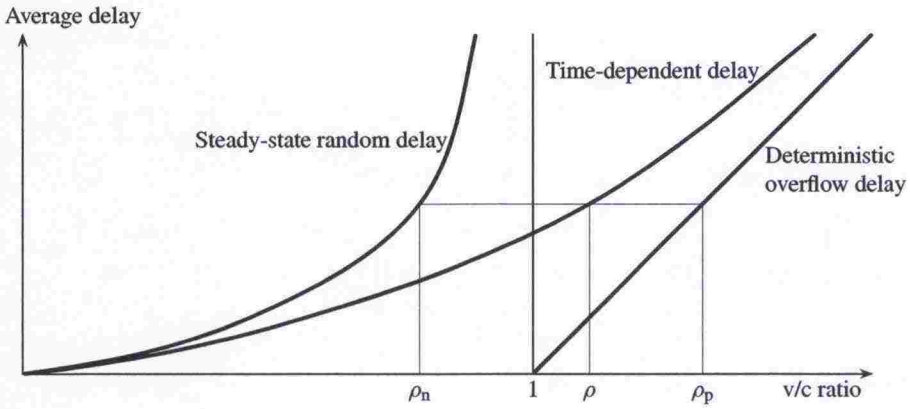


Figure 4.9: Coordinate transformation

The degree of saturation ρ_n, which produces the same average delay in stochastic steady state analysis as ρ in the stochastic time-dependent analysis, is

$$\rho_n = \rho - (\rho_p - 1) = \rho - \frac{2w_o(t)}{t}. \tag{4.63}$$

Because the average deterministic overflow delay $w_o(t)$ at ρ_p is equal to the time-dependent delay w_r at ρ as well as the stochastic steady state delay w_D at ρ_n, the time-dependent random delay component can be written as

$$w_r = \frac{\alpha\rho_n}{C(1-\rho_n)} = \frac{\alpha\left(\rho - \frac{2w_r}{t}\right)}{C\left(1 - \rho + \frac{2w_r}{t}\right)}. \tag{4.64}$$

(The parameter t is not displayed with w_r in order to keep the symbols simple.) We obtain

$$\frac{2}{t}w_r^2 - \left(\rho - 1 - \frac{2\alpha}{Ct}\right)w_r - \frac{\alpha\rho}{C} = 0. \tag{4.65}$$

The average time-dependent delay is (cf. Catling 1977, Kimber & Hollis 1978, Burrow 1989)

$$w_r = \frac{t}{4}\left(\rho - 1 - \frac{2\alpha}{Ct} + \sqrt{\left(\rho - 1 - \frac{2\alpha}{Ct}\right)^2 + \frac{8\alpha\rho}{Ct}}\right). \tag{4.66}$$

The negative branch of the solution has no physical significance (Kimber & Hollis 1979). In the M/D/1 queuing model $\alpha = 1/2$, and the equation becomes

$$w_r = \frac{t}{4}\left(\rho - 1 - \frac{1}{Ct} + \sqrt{\left(\rho - 1 - \frac{1}{Ct}\right)^2 + \frac{4\rho}{Ct}}\right). \tag{4.67}$$

Figure 4.10 displays the steady state random delay, time-dependent delay and deterministic overflow delay for an overflow period generated by a five-minute peak period.

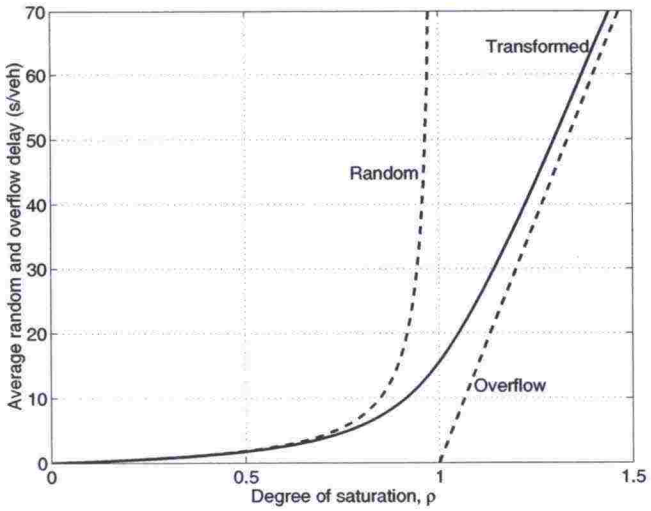


Figure 4.10: Steady-state random delay, transformed time-dependent delay and deterministic overflow delay of an overflow period generated by a five-minute peak period, when cycle length is 100 s, green time 50 s, and capacity 1,000 veh/h

The total average control delay is

$$w = w_u + w_r. \tag{4.68}$$

Catling (1977) used equation $w = w_0 + w_r$, where $w_0 = 0.5c(1 - g/c)^2$ is the average delay under very low flow conditions (see also Shawaly, Ashworth & Laurence 1988). w_0 is obtained from equation (4.13) by setting $\rho = 0$. This approach does not, however, consider the increase in uniform delay due to high degrees of saturation.

As Hurdle (1984) remarks, the coordinate transform method is not a result of any detailed analysis of queue behavior. The only justifications for the method is that it provides a smooth transition from steady-state analysis to time-dependent analysis in a way that satisfies the intuitive ideas of what ought to happen. Also, the method does not consider the length of the overflow-queue discharge process following the peak period.

Akçelik (1988) has generalized the 1985 HCM (Transportation Research Board 1985), Australian (Akçelik 1981) and Canadian (Teply, Allingham, Richardson & Stephenson 1995) formulas for time-dependent delay as

$$w_r = 900t\rho^{a_1}\left(\rho - 1 + \sqrt{(\rho - 1)^2 + \frac{a_2(\rho - \rho_0)}{Ct}}\right), \tag{4.69}$$

where t is the length of the analysis period (in hours), a_1 and a_2 are calibration parameters, and ρ_0 is the degree of saturation below which the average overflow queue is

approximately zero. The $2\alpha/(Ct)$ term in equation (4.65) is ignored. Burrow (1989) rewrote the equation in a more general form

$$w_r = 900t\rho^n \left(\rho - 1 + b_1 + \sqrt{(\rho - 1)^2 + \frac{m(\rho + b_2)}{Ct}} \right), \tag{4.70}$$

where b_1 and b_2 are additional parameters consistent with equation (4.66).

For a more extensive overview of the coordinate transformation method see Rouphail, Tarko & Li (1997).

4.6 Delay in HCM2000

The service measure in HCM2000 (Transportation Research Board 2000) is control delay (Table 4.1). The average control delay per vehicle is

$$w = w_u f_p + w_r + w_q, \tag{4.71}$$

where w_u is the uniform control delay (4.16), f_p is uniform delay progression adjustment factor, w_r is incremental delay (4.73), and w_q is initial queue delay. This approach has been called the *generalized delay model*. The method was essentially presented already in the 1997 HCM (Transportation Research Board 1998). HCM2000 introduced only minor modifications.

The progression adjustment factor is

$$f_p = \frac{(1 - P_g) f_s}{1 - \frac{g}{c}}, \tag{4.72}$$

where P_g is the proportion of vehicles arriving during green phase, and f_s is supplemental adjustment factor for platoon arriving during green. The value of f_p can be obtained from Exhibits 16–11 and 16–12 in HCM2000 based on arrival type and green ratio. Six arrival types (Exhibit 16-4 in HCM2000) describe platoon progression from very poor progression quality (AT1) to exceptionally good progression quality (AT6). Table 4.2 provides the HCM2000 default values for the supplemental adjustment factor f_s for platoon arriving during green as well as for the platoon ratio $R_p = P_g c/g$. If field data is not available, the proportion can be estimated as $P_g = R_p g/c$, and f_p is obtained from equation (4.72).

Table 4.2: Default values for platoon ratio R_p and supplemental adjustment factor f_P in HCM2000

Arrival type	Platoon ratio R_p	Supplemental adjustment f_s
1	0.333	1.00
2	0.667	0.93
3	1.000	1.00
4	1.333	1.15
5	1.667	1.00
6	2.000	1.00

The HCM2000 (Transportation Research Board 2000) ignores the $2\alpha/(Ct)$ terms in equation (4.66), expresses the analysis period T in hours and capacity C in veh/h, and divides α into two separate adjustment factors k and I. The average *incremental delay* is

$$w_r = 900T \left(\rho - 1 + \sqrt{(\rho - 1)^2 + \frac{8kI\rho}{CT}} \right) \quad \text{(s/veh)}, \tag{4.73}$$

where k is the incremental delay factor that is dependent on controller settings, and I is the upstream filtering/metering adjustment factor. For isolated pretimed signals $k = 0.5$ and $I = 1.0$ resulting in equation (4.67), except for the terms $(Ct)^{-1}$. The HCM2000 incremental delay is slightly higher (Fig. 4.11).

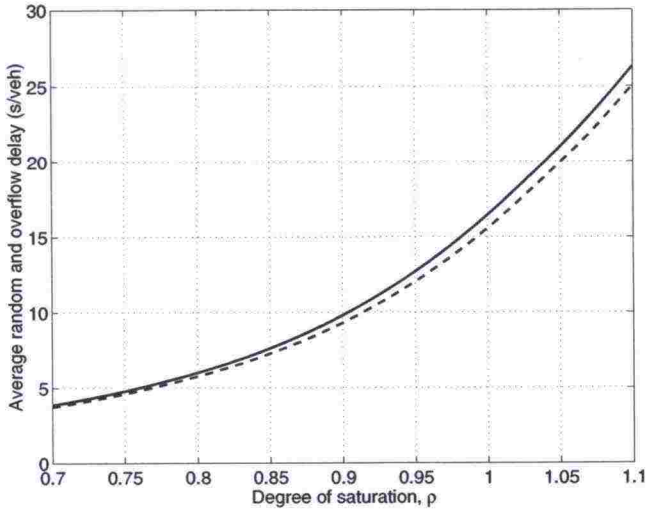


Figure 4.11: HCM2000 incremental delay (solid curve) compared with the time-dependent random delay (dashed curve) of equation (4.67) and Figure 4.10

The delays of fixed control are adjusted for traffic-actuated control with calibration factor k, which has been derived from queuing theory. The value of k-parameter, ranging from 0.04 to 0.50 (Exhibit 16–13 in HCM2000), depends on the degree of saturation and the unit extension interval of the traffic-actuated control. With pretimed control, random arrival process and constant departures, the value of k is 0.50. With high degrees of saturation, the fixed control and VA-control tend to behave in a similar manner and the k-parameter converges to 0.50. As unit extension and degree of saturation decrease, k decreases also, reaching the minimum value of 0.04 when unit extension is ≤ 2.0 s and $\rho \leq 0.50$.

The calibration factor k has an effect on the incremental (random and overflow) delay only (see Fig. 4.12), not on the uniform delay component. The effect of k on the incremental delay is largest at low degrees of saturation, but at these conditions the effect of incremental delay on the total control delay becomes negligible. Consequently, it is important that cycle lengths and phase lengths used in the analysis approximate the actual average cycle lengths and phase lengths as closely as possible. Design parameters cannot be used in the analysis, as Figure 4.12 clearly illustrates.

The initial queue delay accounts for delay to all vehicles in analysis period due to initial queue at the beginning of the analysis period. It enables the estimation of delay, when oversaturation extends over multiple analysis periods. The estimation of w_q is described in Appendix F of Chapter 16 in HCM2000.

For pedestrians at signalized intersections the service measure is the average delay per pedestrian on a crosswalk. It is given by equation

$$w_p = \frac{0.5(c-g)^2}{c}. \tag{4.74}$$

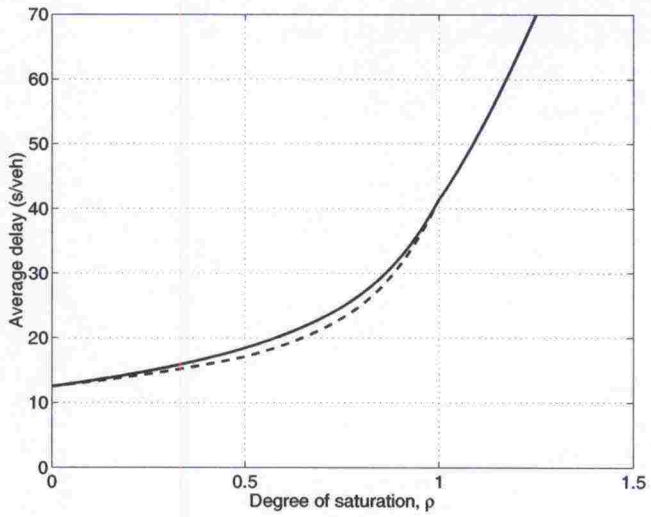


Figure 4.12: Delay in pretimed (solid curve) and traffic-actuated (dashed curve) isolated intersections according to HCM2000 using the same parameters as in Figure 4.10 and a three-second unit extension

As the delay increases the likelihood of noncompliance (i.e., disregard for signal indications) increases. Table 4.3 displays the LOS thresholds and a guide for the likelihood of pedestrian noncompliance.

Table 4.3: Level-of-service criteria and the likelihood of noncompliance for pedestrians at signalized intersections according to HCM2000 (Transportation Research Board 2000)

LOS	Delay (s/ped)	Likelihood of noncompliancs
A	< 10	Low
B	10–20	
C	20–30	Moderate
D	30–40	
E	40–60	High
F	> 60	Very high

Rouphail, Anwar, Fambro, Sloup & Perez (1997) compared the generalized delay model of HCM2000 with TRAF-NETSIM delays and field studies for traffic-actuated traffic signals. They concluded that

1. The generalized delay model and NETSIM yielded comparable delays for basic traffic-actuated control.

2. The generalized model delay was comparable to observed delay in the field.

3. The proposed generalized delay model is sensitive to changes in traffic volumes and traffic-actuated controller settings.

4. The generalized model is an improvement over the then current HCM (Transportation Research Board 1994) model for estimating delay at traffic-actuated traffic signals.

Figures 2–4 in the paper, however, suggest that large delays may be higher than the delays obtained by simulation and field studies. On the other hand, small delays appear to be lower than field measured delays.

Another study (The University of Florida Transportation Research Center 1996) observed that the generalized delay model produced slightly higher delay estimates than NETSIM. The definition of parameter k in this study was different than in the HCM2000.

The delays at oversaturated conditions were found to be in close agreement with those simulated by TRAF-NETSIM. On average, simulated delays were overestimated by about four seconds, but this error was considered small compared with actual delays (Englebrecht, Fambro, Rouphail & Barkawi 1997).

The generalized model can applied to all degrees of saturation. It is not restricted to degrees of saturation less than 1.2, as it was previously (Transportation Research Board 1994). However, the model cannot deal with the interaction between intersections under oversaturated conditions. Also, an analyst should consider the effect of oversaturation on traffic demand.

Taale & van Zuylen (2001) compared the 1997 HCM (Transportation Research Board 1998) method with simulation studies and field measurements. The 1997 HCM method is essentially identical with the HCM2000 method. The HCM outcome was very good when compared with simulated pretimed control and simple traffic-actuated control. The results were much worse for real-life advanced traffic-responsive control (RWS C-controller), especially at degrees of saturation above 0.8. The variation patterns that occur in real-life were considered as one explanation for these results. Another source of error was the blocking effect of short turning lanes.

HCM overestimated short delays and underestimated long delays for simple traffic-actuated control. The reason for this appeared to be the averaging of cycle lengths and green intervals over the total observation period. HCM delay estimates were quite realistic when the analysis was performed for short time periods using observed cycle lengths and green intervals as parameters. For traffic-actuated control the agreement between the HCM method and simulation results depended strongly on the accuracy with which the average cycle length and green interval were known for a given time period. The authors also observed a considerable variation in saturation flows.

4.7 Delay in Capcal 2

Capcal has been the major capacity analysis software used in Finland for intersections. Capcal was developed in Sweden as an implementation of the Swedish capacity manual (Statens vägverk 1977). The software as well as the user manual were translated into Finnish by the Finnish Roads and Waterways Administration (Tiensuunnittelutoimisto 1987). The second version of the software, Capcal 2 (Vägverket 1995), is an enhanced version of the software introducing new analysis methodologies.

In Capcal 2 (Swedish National Road Administration 1995) the total average delay (*medelfördröjning totalt*) is the sum of interaction delay (*medelfördröjning stopplinje*) and geometric delay (*medelfördröjning övrigt*). Because the geometric delay component includes the delay due to deceleration and acceleration, the interaction delay can be interpreted as the stop delay. However, it is assumed that at least some part of the deceleration component should be neglected, because deceleration to some extent takes place when a vehicle catches up a queue.

For stopping vehicles the additional delay due to intersection geometry is less than for vehicles with little or no control delay. For simplicity, the total average delay w

includes the average acceleration delay w_a and the larger value of average interaction (stopped) delay w_s and average deceleration delay w_d:

$$w = (\max\{w_s, w_d\} + w_a) f_{va}, \tag{4.75}$$

where f_{va} is the correction factor for vehicle-actuated control. For pretimed control $f_{va} = 1$.

The delay definitions of Capcal 2 are not compatible with the HCM delay definitions. In HCM2000 the control delay includes the deceleration and acceleration delay due to traffic signals and vehicle interaction. The control delay does not, however, include the delay due to the geometry of the intersection. Following the approach adopted in the 1985 HCM (Transportation Research Board 1985) it can be approximated that the interaction delay (stop delay) is 77 % of control delay (cf. Akçelik 1988). The total delay in Capcal 2 is, however, larger than the control delay. *In the calculations below the interaction (stop) delay of Capcal 2 has been used.*

The interaction (stop) delay is

$$w_s = w_u + w_r, \tag{4.76}$$

where w_u is the uniform delay (4.13) and w_r is the random delay. For degree of saturation $\rho \leq 0.8$ the random delay follows the M/D/1 model (4.41); i.e., $w_r = w_D$, so that the "theoretical" average delay follows the unadjusted Webster model (4.43). The results are approximately 10 % higher than the Webster delay (cf. 4.45). For degrees of saturation $\rho > 1.4$ the deterministic overflow delay model (4.33) is used; i.e., $w_r = w_o$. Capcal 2 does not use the coordinate transformation method, but for $0.8 < \rho \leq 1.4$ the random component is estimated by interpolation (Fig. 4.13). The effects of blocked short lanes are estimated with a macro-simulation procedure.

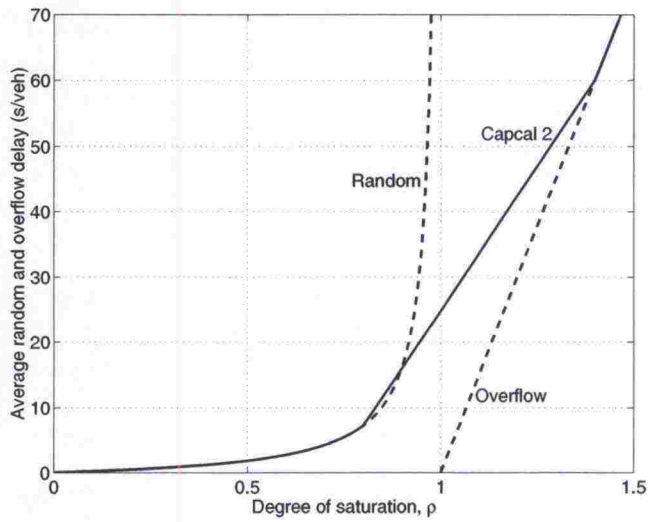


Figure 4.13: Steady-state random delay, deterministic overflow delay, and the Capcal 2 time-dependent random delay (solid line)

Capcal 2 has two models for traffic-responsive control. The "older" type (type 1) has a fixed phase sequence. Vehicles are detected for each signal group. The "modern" type (type 2) control assumes a signal-group controller with pedestrian push-buttons, all-red rest state etc. Arrivals are detected per lane.

Type 2 control has lower delays at low volumes. As the traffic volume approaches null, the probability of an arrival during an all-red rest state increases, and the average delay converges to zero. This is a major difference between HCM2000 (Fig. 4.12) and Capcal 2. The latter method can be expected to be a better model for Finnish traffic-responsive control, especially at low flow conditions.

The correction factor for modern traffic-actuated control is

$$f_{va} = (1 - P_r) + P_r \min\{1, C_2\}, \tag{4.77}$$

where P_r is the probability that the signal is all red when the traffic flow is zero, when also pedestrian flow is considered. The probability is calculated as

$$P_r = \prod_{1}^{N_A} \left(1 - e^{q_{ped} \sum g_{min}}\right), \tag{4.78}$$

where N_A is the number of approaches, $\sum g_{min}$ is the sum of minimum green times for all lanes in the approach, and q_{ped} is the pedestrian flow. The second parameter C_2 is calculated as

$$C_2 = 0.1 + \rho(1.9 - C_1) - \rho^2(1 - C_1), \tag{4.79}$$

where

$$C_1 = 0.1 \left[\min\{4, N_P - 2\} - 0.4 \left(1 - \frac{\rho}{\rho_{max}}\right)\right], \tag{4.80}$$

ρ is the degree of saturation in the actual lane, ρ_{max} is the maximum degree of saturation for any lane, and N_P is the number of phases. Contrary to the approach in HCM2000, Capcal 2 adjusts the total delay (including the uniform delay component) for vehicle actuation. Capcal 2 is also easier to use, because it is based on the correction factor applied to delays due to pretimed control.

4.8 Delay in DanKap

DanKap is an implementation of the new Danish capacity methodology (Vejdirektoratet 1999*b*, Vejdirektoratet 1999*a*). The delay for signalized intersections is estimated as

$$w = w_u f_p + w_r, \tag{4.81}$$

where w_u is the uniform delay component (4.16), f_p is arrival type adjustment factor (4.72), and w_r is incremental delay

$$w_r = \frac{t}{4} \left(\rho - 1 + \sqrt{(\rho - 1)^2 + \frac{4\rho}{Ct}}\right), \tag{4.82}$$

where Ct is the maximum number of departing vehicles during the analysis period t. Time is expressed in seconds.

In fact, DanKap provides the HCM2000 control delay formula (4.73) for pretimed isolated intersections; i.e., $k = 0.5$ and $I = 1.0$. The arrival type adjustment factor f_p has similar values as presented in HCM2000 Exhibit 16-12, but expressed with less significant digits. DanKap has no adjustment for traffic-responsive control (k) or upstream filtering/metering (I).

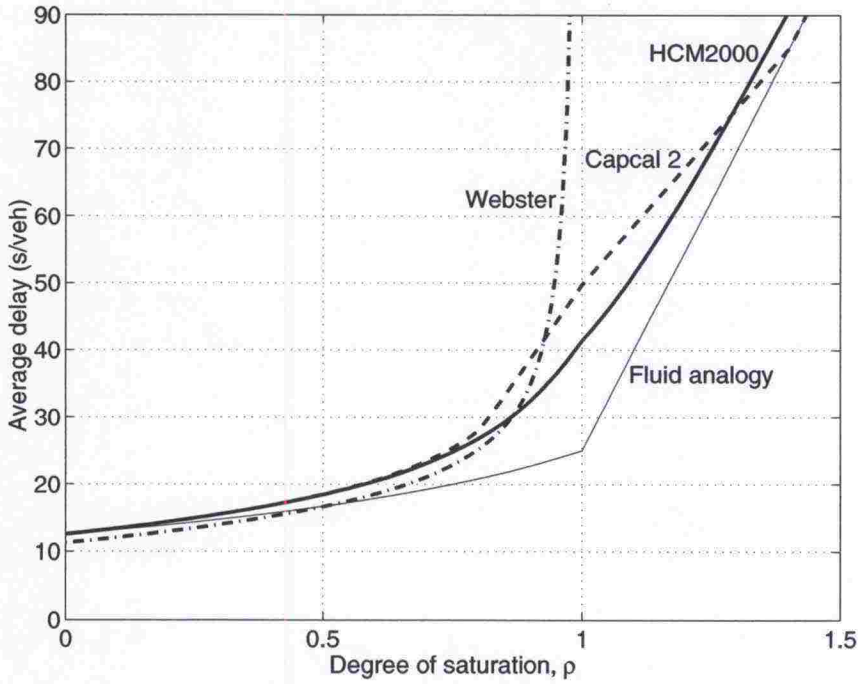


Figure 4.14: Delay in pretimed isolated intersections according to the fluid analogy model, shorter Webster equation (4.45), Capcal 2, and HCM2000 using the same parameters as in Figure 4.10

Figure 4.14 displays a comparison of various delays models for pretimed isolated intersections using the same parameters as in Figure 4.10. Under these conditions DanKap gives same results as HCM2000. Observation period is five minutes. The parameters have been chosen so that the most important properties and differences in the models are visible.

At low degrees of saturation all models give similar results. The Webster delays are, however, 10 % lower because of the 0.9 factor in equation (4.45). At moderate degrees of saturation Capcal 2 and HCM2000 delays are slightly higher than the Webster delays. As the degree of saturation approaches unity, Webster delays increases rapidly, but Capcal 2 and HCM2000 delays adjust the delay for the limited length of the overflow period. For degrees of saturation below 0.8 the coordinate transformation method of HCM2000 gives delays slightly lower than Capcal 2. When the degree of saturation is close to unity, the interpolation method of Capcal 2 gives higher delays, but HCM2000 delays become higher as the degree of saturation approaches and exceeds 1.4. At $\rho = 1$ there is an inflection point in the Capcal 2 and HCM2000 curves, because the uniform delay component (4.16) increases with increasing degree of saturation as long as $\rho < 1$, but becomes constant ($w_u = r/2$) at oversaturated conditions (see Fig. 4.2).

4.9 Delay estimation in Finland

The Finnish signal control handbook (Kehittämiskeskus 1996) use two indicators for the quality of service: operational quality ('toimivuus', Table 4.4) and level of service (Table 4.1), with no preference. The major Finnish traffic and transportation engineering handbook (Lyly 1988) presents only the operational quality, whereas the Capcal

software (Tiensuunnittelutoimisto 1987) prints only the level of service.

Table 4.4: Operational quality in Finnish signalized intersections (Kehittämiskeskus 1996)

Degree of saturation	Utilization factor	Operational quality	Congestion
< 0.85	< 0.9	Good	No congestion
0.85...0.95	0.9...1.0	Satisfactory	Occasional congestion
0.95...1.05	1.0...1.1	Tolerable	Short-term congestion and queues
> 1.05	> 1.1	Bad	Long-lasting congestion and queues

Operational quality has four classes: good, satisfactory, tolerable and bad (see table 4.4). The measures of effectiveness are degree of saturation and utilization factor. The concept was introduced in an older handbook (Pohjoismaiden tieteknillinen liitto 1978), where utilization factor was the performance measure.

The degree of saturation (critical v/c ratio) is

$$\rho = \frac{\sum_{j=1}^{n} y_j}{1 - \frac{\sum_{j=1}^{n} L_j}{c}} = \frac{\sum_{j=1}^{n} \max_{i \in m} \left\{ \mu_j(i) \left(\frac{q_i}{s_i} \right) \right\}}{1 - \frac{\sum_{j=1}^{n} L_j}{c}} \tag{4.83}$$

where

ρ = degree of saturation for the intersection
m = number of lane groups
n = number of phases
y_j = maximum (critical) flow ratio of phase j
s_i = saturation flow of lane group i
q_i = demand flow rate of lane group i
c = cycle length
L_j = total lost time preceding phase j.

The membership function $\mu_j(i)$ is defined as

$$\mu_j(i) = \begin{cases} 1, & \text{if lane group } i \text{ has green in phase } j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.84}$$

If a lane group can have green in several consecutive phases, the flow ratio is calculated for the the whole green length, and it is compared against the sums of flow ratios of other conflicting lane group combinations during these phases. The green signal in two consecutive phases is, of course, not interrupted by yellow change or red clearance intervals.

The utilization factor indicates the proportion of cycle length utilized by lost times and arriving vehicles in the critical signal groups:

$$\rho' = \sum_{j=1}^{n} y_j + \frac{\sum_{j=1}^{n} L_j}{c} = \sum_{j=1}^{n} \max_{i \in m} \left\{ \mu_j(i) \left(\frac{q_i}{s_i} \right) \right\} + \frac{\sum_{j=1}^{n} L_j}{c}. \tag{4.85}$$

Degree of saturation and utilization factor use a different approach to evaluate the effect of lost times in the cycle. At low degrees of saturation the utilization factor is higher than the degree of saturation. At high degrees of saturation it is lower.

The level of service in the Finnish handbook has the same criteria (Table 4.1) as HCM 1985 (Transportation Research Board 1985). The type of average delay is not

explicitly specified, but it can be assumed that stop delay is intended, as in the 1985 HCM. The most commonly used delay model, the Webster model (4.45), however, gives control delays. The practitioners should be very careful and use a delay model compatible with the critical values.

The Capcal software used in Finland reports control delays, but uses stop delays to estimate the level of service (Tiensuunnittelutoimisto 1987). The control delays are calculated following the Swedish method (Statens vägverk 1977). It is an implementation of the Webster method (4.46) with a slight modification:

$$w = 0.9\left(Ac + \frac{B}{c}\right), \tag{4.86}$$

where the Webster table for B is presented, and the value for A is obtained from a chart.

Average delay is considered insufficient as a service measure, because it does not include the effect of stops. Under low volume conditions the average delays may be short, but a considerable proportion of vehicles may have to stop. Especially on major highways the number of stops should be considered more important than delays (Kehittämiskeskus 1996). However, the manual does not present a similar discussion about the utilization factor, which the same problem with the number of stops. It is possible to have a high average delay and a high proportion of stopped vehicles under a low degree of saturation due to a poor signal coordination. Because delay as a service measure considers both degree of saturation and progression quality, it can be preferred over the utilization factor.

Mäkelä (1997) has suggested the application of the German (Brilon, Großmann & Blanke 1994) LOS criteria for both vehicles (Table 4.5) and pedestrians (Table 4.6). Isolated and coordinated intersections have separate LOS criteria. The service measures for vehicles are average control delay and the degree of saturation. The German delay criteria for isolated intersections are higher than in the HCM2000 (Table 4.1). Coordinated intersections have lower criteria for uncongested conditions. For levels of service E and F the criteria are equal. The service measure for pedestrians at isolated intersections is the maximum delay. At coordinated signals the delay criteria are higher, and the average number of stops is used as an auxiliary service measure.

Table 4.5: Level-of-service criteria for vehicles at isolated and coordinated signalized intersections in Germany (Brilon et al. 1994)

	Isolated		Coordinated	
LOS	Average delay (s)	Degree of saturation	Average delay (s)	Degree of saturation
A	$\leq$ 25	—	$\leq$ 5	—
B	$\leq$ 40	—	$\leq$ 15	—
C	$\leq$ 60	—	$\leq$ 40	—
D	$\leq$ 80	$\leq$ 0.85	$\leq$ 60	$\leq$ 0.85
E	$\leq$ 100	$\leq$ 1.00	$\leq$ 100	$\leq$ 1.00
F	> 100	> 1.00	> 100	> 1.00

Table 4.6: Level-of-service criteria for pedestrians at isolated and coordinated signalized intersections in Germany (Brilon et al. 1994)

	Isolated	Coordinated	
LOS	Max delay (s)	Max delay (s)	Number of stops (1/ped)
A	≤ 30	≤ 40	≤ 0.5
B	≤ 40	≤ 50	≤ 1.0
C	≤ 50	≤ 70	≤ 1.5
D	≤ 60	≤ 90	≤ 2.0
E	≤ 70	≤ 110	> 2.0
F	> 70	> 110	—

5 SATURATION FLOWS AND CAPACITY

Saturation flow rate (s) of a lane or a lane group is the flow rate of departing passenger cars at stop line under saturated conditions assuming that the signal is green at all times (cf. Teply & Jones 1991). It is the inverse of the average discharge headway (h_d) at stop line from a continuous queue of passenger cars. The capacity C of a lane or a lane group is

$$C = s\frac{g}{c}, \tag{5.1}$$

where g is the effective green time of the lane or lane group, and c is the cycle length. Saturation flow rate is a basic measure in the estimation of capacity and delay, as the HCM2000 methodology (Fig. 5.1) indicates.

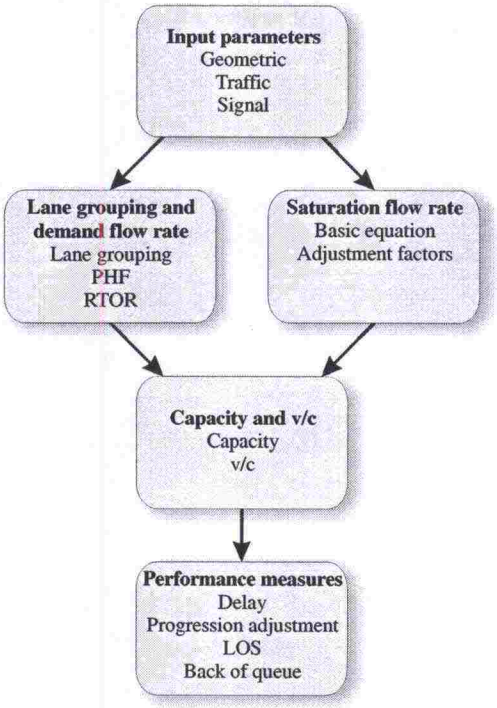


Figure 5.1: Signalized intersection methodology in HCM2000 (Transportation Research Board 2000)

For a lane with different movements the saturation flow rate is the inverse of the average headway weighted by the proportions (p_i) of the flow rates of respective movements:

$$s = \frac{1}{h_d} = \frac{1}{\sum_i p_i h_i} = \frac{1}{\sum_i \frac{p_i}{s_i}}, \tag{5.2}$$

where h_i is the average discharge headway of movement i.

Saturation flow rate is low in the beginning of the green period due to the acceleration delays, and during the yellow change interval, when vehicles decelerate and stop (see Fig. 2.7 on page 29). It is usually assumed that the saturation flow rate first increases and then reaches a constant level, which is maintained during the rest of the green period. There is some evidence that the saturation flow rate increases during the first ten green seconds. During a long green period the saturation flow rate stays at this high level for

about 30 seconds, after which it starts to decrease (Wildermuth 1962, Teply 1981, Teply & Jones 1991, Swedish National Road Administration 1995).

If the effective green time is adjusted for the acceleration and deceleration effects only, the average saturation flow rate may be a function of the effective green length. Teply (1981) has also observed an increase in the saturation flow rate at the end of green and during the initial yellow change interval. He suggested that this may be caused by frequent violations of all-red period or/and by the cases when a large vehicle with a high passenger car unit equivalent enters the intersection as the last vehicle.

5.1 Saturation flow rates in HCM2000

In HCM2000 (Transportation Research Board 2000) the saturation flow rate for each lane or lane group is determined using to the following equation:

$$s = s_0 N f_w f_{HV} f_g f_P f_{bb} f_a f_{LU} f_{LT} f_{RT} f_{Lpb} f_{Rpb} \tag{5.3}$$

where

s = saturation flow rate for subject lane group (veh/h)
s_0 = base saturation flow rate per lane (pc/h/lane)
N = number of lanes in the lane group
f_w = adjustment factor for lane width
f_{HV} = adjustment factor for heavy vehicles
f_g = adjustment factor for approach grade
f_P = adjustment factor for parking
f_{bb} = adjustment factor for stopping buses
f_a = adjustment factor for area type
f_{LU} = adjustment factor for lane utilisation
f_{LT} = adjustment factor for left turns in the lane group
f_{RT} = adjustment factor for right turns in the lane group
f_{Lpb} = pedestrian adjustment factor for left-turn movements
f_{Rpb} = pedestrian adjustment factor for right-turn movements.

The base saturation flow rate is usually 1900 pc/h/lane. In CBD it is ten percent lower $f_a = 0.9$.

The adjustment factor for lane width is

$$f_w = 1 + \frac{w_l - 3.6}{9}, \tag{5.4}$$

where w_l is the lane width in meters. Heavy vehicles and grade have separate adjustment factors:

$$f_{HV} = \frac{100}{100 + P_{HV}(E_{HV} - 1)} \tag{5.5}$$

$$f_g = 1 - \frac{\gamma}{200}, \tag{5.6}$$

where P_{HV} is the percent heavy vehicles, E_{HV} is the passenger car equivalency ($E_{HV} = 2$) pc/HV, and γ is the grade (%) on a lane group approach. Separate adjustment factors for heavy vehicles and grade indicate that grade has an effect on the performance of passenger cars also.

The adjustment factors for turning traffic are perhaps the most complex and significant factors in the estimation the saturation flow rate. The adjustments are based on several parameters:

1. Type of lane: exclusive or shared
2. Type of signal phasing: permitted or protected
3. Conflicting pedestrian and vehicle flow volumes
4. Proportion of turning traffic.

The adjustment factor (f_{RT}) for right turning traffic depends on lane type and the proportion of right turning vehicles: 0.85 for exclusive lanes, $1.0 - 0.15P_{RT}$ for shared lanes, and $1.0 - 0.135P_{RT}$ for a single lane, where P_{RT} represents the proportion of right turners. Factor f_{Rpb} adjusts for the blocking effect of pedestrians and bicycles.

The intersection geometry has a lower impact on left turn adjustment, while traffic flow conditions are more important. The the left-turn adjustment is particularly significant in the case of permitted left-turn movements on shared lanes, when the opposing traffic flow blocks the left turning traffic. In addition to left turners, the through driving traffic flow is disturbed. For protected left turns, the adjustment factor is 0.95 for exclusive lanes and $1/(1 - 0.05P_{LT})$ for shared lanes, where P_{LT} represents the proportion of left turners.

The impact of permitted left turns is complex. The adjustment for permitted left turns in HCM2000 depends on the lane type (exclusive/shared) and the degree of saturation of through and left turning traffic, both on the studied approach and the opposing flow. The estimation of adjustment factors for permitted phases with leading or lagging protected phases is performed by dividing the green time into separate time intervals based on control (protected/permitted) and queue discharge (discharging/discharged). The special procedures for nonprotected left turns are given in Appendix C of Chapter 16 in the manual.

5.2 Saturation flow rates in DanKap

While DanKap (Vejdirektoratet 1999*b*) follows the HCM2000 delay estimation methodology, the procedure for estimation of saturation flow rates is much simpler and adjusted to the Danish conditions. DanKap calculates the basic saturation flow rate as

$$s = \frac{3600 f_{LT}}{h_d}, \tag{5.7}$$

where f_{LT} is an adjustment factor for permitted left turns, and h_d is the average discharge headway (Table 5.1). The "permitted right turn" in table 5.1 indicates a right-turn movement with bicycle and/or pedestrian interference.

*Table 5.1: Average discharge headways and saturation flows in simple traffic movements according to DanKap (Vejdirektoratet 1999*b*)*

Movement	Discharge headway (s/pc)	Saturation flow (pc/h)
Left turn	2.2	1,636
Through	1.8	2,000
Protected right turn	2.2	1,636
Permitted right turn	3.2	1,130

The left turn adjustment factor (f_{LT}) is defined as

$$f_{LT} = \frac{h_d q_o e^{-5.5 q_o/3600}}{3600\left(1 - e^{-3.0 q_o/3600}\right)}, \tag{5.8}$$

where q_o is the opposing flow rate (pc/h). The saturation flow rate for a basic permitted left turn is then

$$s_{LT} = \frac{q_o e^{-5.5 q_o/3600}}{\left(1 - e^{-3.0 q_o/3600}\right)}. \tag{5.9}$$

Figure 5.2 displays the basic saturation flow rate of permitted left turns as a function of opposing flow rate. When the opposing flow approaches null, the saturation flow rate approaches 1,200 pc/h, which is well below 1,636 pc/h, the saturation flow rate of protected left and right turns.—DanKap presents traffic volumes per observation period (T). The equations have been transformed to hourly flow rates.

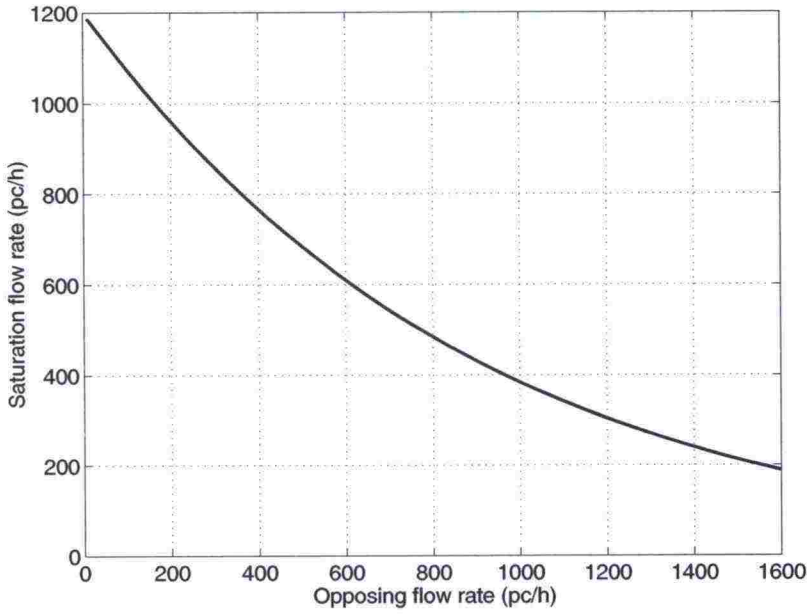


Figure 5.2: Basic saturation flow rate for permitted left turns according to DanKap (Vejdirektoratet 1999b)

Capacity is

$$C = s \frac{g}{c} f_{HV}, \tag{5.10}$$

where the effective green time (g) is the actual green time plus one second, and f_{HV} is the adjustment factor for heavy vehicles. The passenger-car equivalencies (PCEs) are displayed in table 5.2. The adjustment factor is

$$f_{HV} = \frac{\sum_i q_i}{\sum_i q_i E_i}, \tag{5.11}$$

where q_i is the flow rate of vehicle category i, and E_i is its PCE value.

5.3 Saturation flow rates in Capcal 2

Capcal 2 (Swedish National Road Administration 1995) suggests 1,850 veh/h as a base saturation flow rate under reference conditions defined as

- 3.5 m wide lane for through traffic only
- Passenger cars only, no bicycles

Table 5.2: Passenger-car equivalencies in DanKap (Vejdirektoratet 1999b)

Vehicle category	PCE
Motor cycles	0.5
PCs and vans	1.0
Trucks and busses	1.5
Trucks with trailers	2.0

- No interference from bus stops or parked vehicles
- Normal road surface conditions; i.e., no ice or snow
- In all other aspects, "average" conditions for Sweden.

Local adjustment of this reference value should be no more than ±10 %. For prevailing conditions the saturation flow rate is obtained as

$$s = 1850 \prod_i f_i \prod_j f_j, \tag{5.12}$$

where f_i are the adjustment factors for a movement, and f_j are the adjustment factors per lane.

The saturation flow rate is not constant. It decreases systematically after 40 seconds, for green periods which are saturated that long.

For a "normal" mix of heavy vehicles the adjustment factor f_{HV} is

$$f_{\mathrm{HV}} = \begin{cases} \left(1 + 0.5P_{\mathrm{HV}} + 2.5P_{\mathrm{HV}}^2 + 0.1P_{\mathrm{HV}}\gamma\right)^{-1}, & 0.0 \le P_{\mathrm{HV}} \le 0.2 \\ \left(1 + P_{\mathrm{HV}} + 0.1P_{\mathrm{HV}}\gamma\right)^{-1}, & 0.2 < P_{\mathrm{HV}} \le 1.0, \end{cases} \tag{5.13}$$

where P_{HV} is the proportion of heavy vehicles, and γ is is the average uphill slope in % over the section 80 m upstream of the stop line ($0 \le \gamma \le 10$). The adjustment is higher for high proportions of heavy vehicles and with uphill slopes. On a level approach the PCEs of heavy vehicles are 1.5 and 2.0 for $P_{\mathrm{HV}} \le 0.2$ and $P_{\mathrm{HV}} > 0.2$, respectively.

Bicycles mixed in a vehicle lane are not included in the calculations, but the saturation flow of (motor) traffic is adjusted by a factor of

$$f_{\mathrm{BC}} = \begin{cases} \left[1 + 0.3(4.0 - w)p_{\mathrm{BC}}\right]^{-1}, & 2.5 \le w \le 4.0 \\ 1, & 4.0 < w. \end{cases} \tag{5.14}$$

The adjustment increases with lane width (w) and disappears at lanes wider than four meters.

Adjustments per lane include factors for radius (turning vehicles), lane geometry and lane markings, parking, and road surface conditions. No adjustment for visibility is assumed to be necessary in normal cases. For turning movements the adjustment factor is a function of the inner radius (R):

$$0.5 + \left(2 + \frac{1000}{R^3}\right)^{-1}. \tag{5.15}$$

The reduction in the saturation flow rate is about 6 % for a radius of 15 m, and 30 % for a 7 m radius.

Capcal 2 considers secondary conflicts of the following types

- Unprotected left turn (conflict with opposing vehicles)
- Turning traffic and pedestrians (left or right turn)
- Unprotected left turn which also has pedestrian conflicts.

The saturation flow per each movement is described by two parameters:

- Blocked time t_b, which reduces the effective green for the movement
- Average saturation flow s_k for each type of time period k within the green interval.

5.4 Saturation flow rates in Finland

The current Finnish handbook (Kehittämiskeskus 1996) reproduces the saturation flow rates of an older handbook (Pohjoismaiden tieteknillinen liitto 1978), which was based on the Swedish manual (Statens vägverk 1977, Kivelä & Pursula 1982). A summary of new Finnish saturation flow studies (Fig. 5.3) is presented, but the older values are used.

The results in Figure 5.3 are based on measurements in the southern Finland supplemented with simulations (Niittymäki 1998). The field measurements were done by the HCM method.

The base saturation flow rate for through traffic according to the Swedish method and the Finnish guidelines is 1,700 pc/h. A comparison with the international results presented above indicates that this is a low value. Recent Finnish research suggests a base value of 1,940 veh/h, which is similar to the results in HCM2000 and DanKap. The German guidelines (Brilon et al. 1994) use 1,800 pc/h as a base saturation flow rate. In Capcal 2 (Swedish National Road Administration 1995) as well as in Australia (Akçelik 1981, Austroads 1988) the base value is 1,850 pc/h. Recent British research (Welsh 2001) has suggested even as high base saturation flow rate as 2,080 pc/h. It appears that the more recent results the higher are the saturation flow rates. The new Finnish results are well in line with the international research.

The saturation flow rate of turning movements without conflicts is 1,500 pc/h in the old guidelines (Pohjoismaiden tieteknillinen liitto 1978). The new values are 1,800 pc/h for left turning movements and 1,750 pc/h for right and left+right turning traffic with a typical right-turning radius of 12 m. The saturation flow rate is lower when the corner radius is shorter (Fig. 5.3). These values are similar to the Capcal 2 model, but higher than in HCM2000 and DanKap. However, the recent British results (Welsh 2001) are even higher than the Finnish estimates. Figure 5.3 displays adjustments for turning movements with pedestrian and/or opposing traffic conflicts as well as corrections for approach grade and corner radius.

The saturation flow rate model for a shared lane (type B) is linear:

$$s_B = a_0 - a_1 P_{LR}, \tag{5.16}$$

where P_{LR} is the percent of turning vehicles and a_0 and a_1 are model parameters (Table 5.3). The radius is not considered in this model. On a lane having mixed through, and left and right turning traffic there is a discontinuity as the proportion of turning traffic approaches zero.—Equation 5.16 should be used to estimate the saturation flow rate of lane type B, not the values presented in Figure 5.3, where the percent turning vehicles is unspecified.

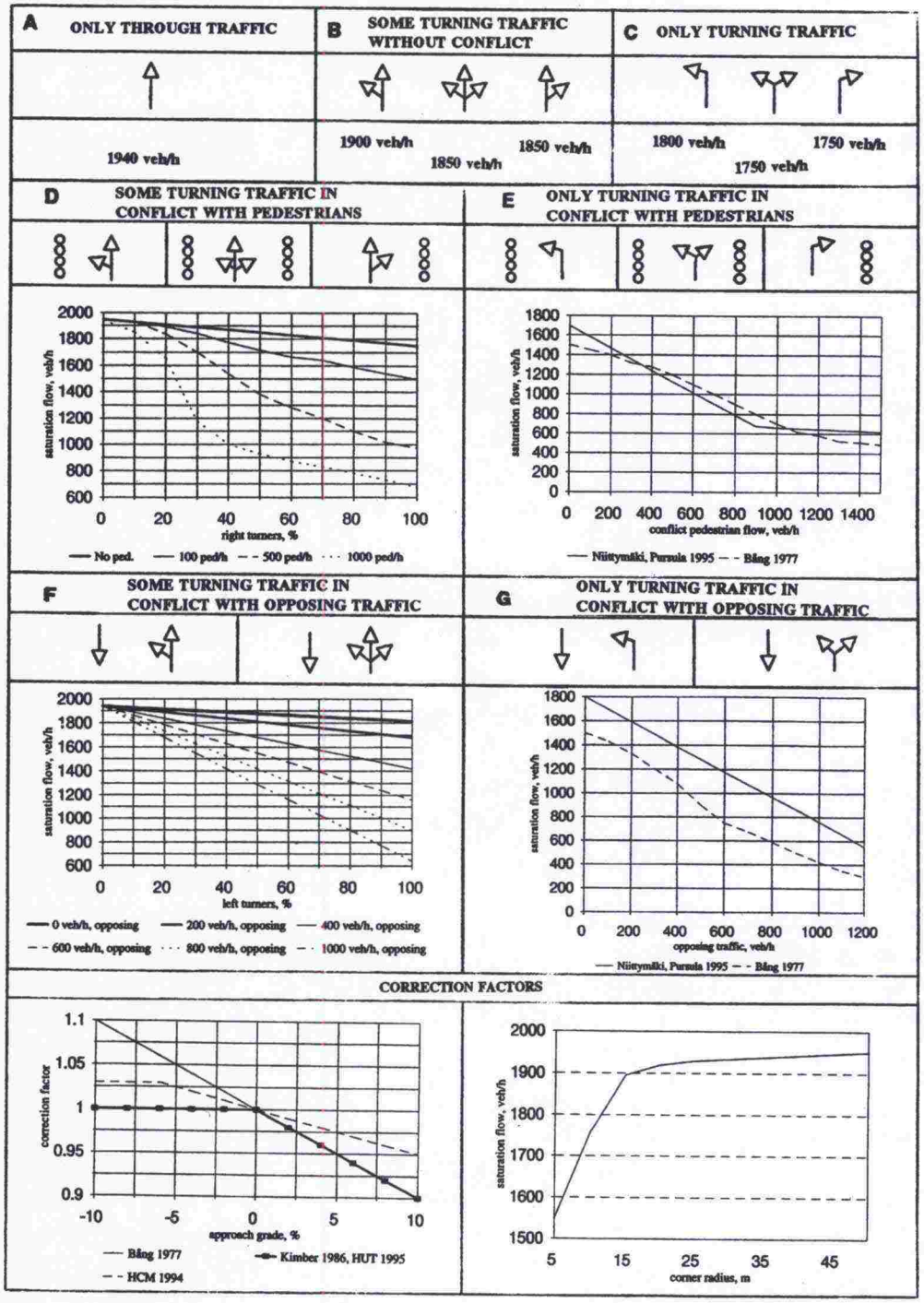


Figure 5.3: Finnish saturation flow guidelines (Niittymäki & Pursula 1997)

Table 5.3: Parameters of saturation flow models for lanes of mixed through and turning traffic according to Niittymäki & Pursula (1997)

Lane type	a_0	a_1
Through + right	1,947	1.96
Through + left	1,946	1.44
Through + left + right	1,925	1.64

The saturation flow rate for a turning lane with pedestrian conflicts (lane type E) is (Niittymäki 1998)

$$s_E = \begin{cases} 1692 - 1.13q_{ped}, & q_{ped} < 900 \\ 660 - 0.083(q_{ped} - 900), & q_{ped} > 900, \end{cases} \quad (5.17)$$

where q_{ped} is the pedestrian flow rate in the pedestrian crossing (Fig. 5.3). A shared left turn lane with permitted turns (type F) has saturation flow rate

$$s_F = 1940 - 0.013q_o P_{LT}, \quad (5.18)$$

where q_o is the opposing flow rate (veh/h) and P_{LT} is the percentage of left turning vehicles. For exclusive left-turn lanes with permitted left turns (lane type G) the saturation flow rate is

$$s_F = 1800 - 1.04q_o. \quad (5.19)$$

A critical gap analysis is suggested, when this method is not considered adequate.

The Finnish method defines saturation flow rates for seven lane types (A–G, Figure 5.3), as in the Swedish capacity manual (Statens vägverk 1977). Accordingly, no adjustment factors are used for the effect of turning movements, opposing traffic and pedestrians, as in HCM2000. This makes the method rather limited for the analysis of combined effects, such as opposing traffic and pedestrians. For lane types (D/F and E/G) having conflicts with opposing traffic and pedestrians, the Swedish manual suggests the lower of the two saturation flow rates (D or F, say). It is, however, possible to convert a saturation flow rate s_i to an adjustment factor $f_i = s_i/1940$.

The adjustment (f_a) for CBD-areas is 0.93. The estimated adjustment factors for road surface and weather conditions are displayed in Table 5.4. Darkness decreases saturation flows by five percent compared to daylight conditions. (Niittymäki & Pursula 1997)

Table 5.4: Adjustment factors for road surface and weather conditions in Finland (Niittymäki & Pursula 1997)

Surface / weather	f_s
Slippery and snowfall	0.75–0.80
Slippery road surface	0.85
Rainy	0.90
Wet road surface	0.95

The results presented are not adequate for the estimation of the heavy-vehicle adjustment. The measured effects were slightly higher than the adjustment factors in the 1994 HCM (Niittymäki & Pursula 1997). The results were, however, very similar to the Capcal 2 model, the use of which can be suggested.

The adjustment factor for approach grade (Fig. 5.3) is $f_g = 1 - \gamma_+$, where γ is the grade. A negative grade (downhill) does not have any effect on saturation flow. Niittymäki & Pursula (1997) report that the loss of saturation flow in uphill grades occurs "mostly because of heavy vehicles". A similar statement is made in the Capcal 2 manual (Swedish National Road Administration 1995). If the heavy vehicle adjustment of Capcal 2 (5.13) is used, no further adjustment for grade is needed. The grade adjustment in Figure 5.3 is similar to the Capcal 2 model for 10 % heavy vehicles. In this case also the Capcal 2 model can be suggested.

6 SIMULATION STUDY

6.1 HUTSIM simulation model

Microscopic simulation of traffic is based on vehicle kinematics, and vehicle-vehicle and vehicle-infrastructure interactions. This allows a detailed analysis of sophisticated traffic control systems.

The simulations for this research were carried out with HUTSIM, which is a Finnish object-oriented microscopic simulation software family. HUTSIM consists of three parts: HUTEDI, HUTSIM and HUTSIM Analyzer. HUTEDI is a model editor operating under the MS-DOS® environment. HUTSIM is a traffic simulator, which also runs under the MS-DOS® environment. HUTSIM Analyzer is a new post processor running under the Microsoft Windows® operating system.

HUTSIM has been under development at the Helsinki University of Technology since 1989. The stochastic properties of approaching traffic and the reactions of conflicting traffic flows can be described and analyzed with sufficient accuracy. HUTSIM has been developed for signal controlled intersections, but it can also be used to analyze unsignalized intersections, small networks, and highway sections. HUTSIM is an application of object-oriented programming and rule based interaction dynamics.

Vehicle dynamics in HUTSIM is based on several models. The most important of which are

1. Vehicle acceleration and deceleration models
2. Car-following model
3. Lane-changing model.

Kosonen (1999) has described each model in detail.

HUTSIM has been extensively calibrated and validated for Finnish signalized intersections (Niittymäki 1993, Niittymäki & Pursula 1994, Niittymäki 1998). The microscopic and macroscopic properties of the system were measured, the HUTSIM parameters were calibrated, and the macro-level results were compared with field measurements. The results of HUTSIM were considered very reliable.

The delays of each vehicle can be calculated easily from the output of the simulations. The movements of each vehicle were stored in a HUTSIM output file. For example, vehicle speeds were updated and stored in an output file every 0.1 seconds, which made the simulation results very informative. The delay calculation was based on the simulated vehicle movements and the desired movements; i.e., driving through the model without any disturbances. The difference between the vehicle movements realized in the simulation and the desired driving behaviour was used as the total delay measure. In addition to the individual delays, the output file included information like vehicle generation times, drivers' desired speeds and speed changes, lane changes, exit times from the model, count of stops, etc. Delays, queue lengths, and percentages of stopping vehicles were calculated for each signal group.

The signal-group oriented control of the HUTSIM simulator represents a decentralized control strategy, in which every single signal group is able to operate according to its own parameters. A signal group updates its status according to time or according to the signals from other control objects, like controller and detector logic. The most important control functions are permission, request and extension of a green signal. These main tasks of the control objects are processed as follows (Kosonen 1996):

1. Controller gives a signal group permission to initiate green (or red). The permission to intiate green is granted, if a signal group has a green request and no conflicting signal group has green extensions. The controller sets also safety-related parameters, such as the minimum green interval for each group, and the intergreen time between two conflict groups.

2. Detectors produce the traffic information that is the basis of green extension decisions. The extensions are granted after the minimum green interval has expired and detection exists. Each detector has a preset extension time. When no unit extension is operating or the controller-set maximum green has been reached, extensions of the green time are terminated, and the signal group switches to red. The group may also remain green, if no demand exists in other approaches (no request for green).

The signal control functions in HUTSIM simulate some, but not all of the minimum functions described above (section 3.2.3). In the test simulations, it was possible to simulate completely the queue discharge function and partly the option zone clearance function. Short red-rest interval prevention and initial green extension were not included in the simulations because of the limitations in the simulation capabilities. However, the vehicle dynamics as well as the most important signal control principles in the simulations can be considered realistic. Consequently, the simulation results can be considered reliable for the comparison of theoretical delay calculation methods and the conditions in Finnish intersections. It would have been possible to use a real signal controller in the simulations, but then the simulations would have run in real time. This would have been too time consuming.

The results of three methods (American HCM2000, Danish DanKap, and Swedish Capcal 2) have been used to compare with the simulated control delays of the HUTSIM software calibrated to Finnish conditions. The delays were estimated for five intersections using both traffic-responsive and pretimed signal control. Five degrees of saturation (0.25, 0.50, 0.75, 0.85 and 0.90) and three traffic conditions (minor/major flows 1:1, 1:2 and 1:4) were included. The degree of saturation was calculated following equation (4.83).

The simulated traffic flows for all minor/major flow combinations were defined by iteration, because the volumes affected both the cycle lengths and the degrees of saturation. The basic saturation flows (s_0) of each lane type (for example through or turning lane, shared or exclusive lane, permitted or protected turn) were estimated according to Figure 5.3. The effect of permitted left-turn movements on the basic saturation flow was approximated with following formula (Niittymäki & Pursula 1997):

$$s_{LT} = s_0 - (0.013 q_o P_{LT}) \tag{6.1}$$

where
s_{LT} = saturation flow rate for a permitted left-turn movement
s_0 = basic saturation flow rate for the lane type
q_o = opposing traffic flow (veh/h)
P_{LT} = percentage of left-turning traffic.

The simulation period was one hour. Simulated traffic included 10 % heavy vehicles. The speed limit of 50 km/h was used in simulations, which is also the default speed limit in Finnish urban areas. The speed distribution used in the simulations allowed some vehicles to drive faster or slower than the speed limit. Pedestrians were not included

in the simulations. Weather conditions were assumed ideal. No parking, bus stops or incidents were included in the simulation.

The pretimed signal control parameters were set according to the Finnish guidelines (Kehittämiskeskus 1996). Cycle lengths and green times were calculated separately for each simulated traffic condition. The cycle lengths were calculated following the Webster equation

$$c = \frac{1.5L + 5}{1 - \sum_{j=1}^{n} y_j} \tag{6.2}$$

where L is the total lost time, and y_j is the critical flow ratio of phase j (Webster 1958, Webster & Cobbe 1966). Green times were allocated by setting critical flow ratios (y_j) for each phase approximately equal. The intergreen time between all conflicting signal groups was six seconds.

The degree of saturation (4.83) is a function of cycle length. If the cycle length is determined using equation (6.2), the degree of saturation can be expressed as

$$\rho = \frac{Y(1.5L + 5)}{L(Y + 0.5) + 5}, \tag{6.3}$$

where Y is the sum of critical flow ratios. Accordingly, the degree of saturation is not a linear function of flow rates, or critical flow ratios, in the intersection (Fig. 6.1).

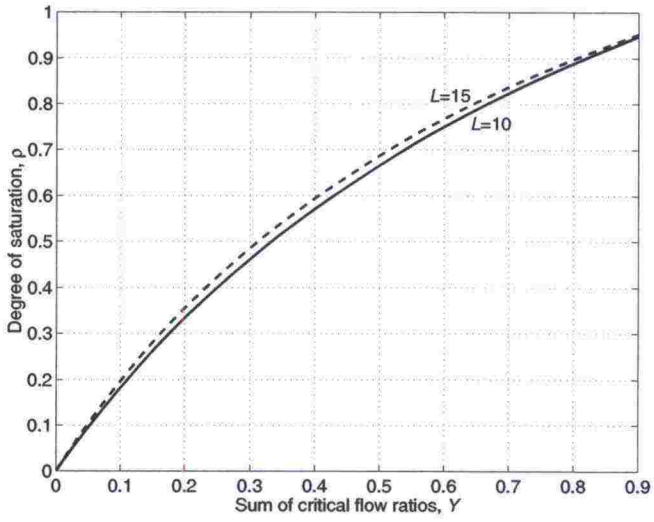


Figure 6.1: Degree of saturation when cycle length is determined using equation (6.2)

Traffic-responsive control had three detectors on each approach/lane. The unit extension time for each detector, green phase starting mode (on demand) and other control parameters were defined in the same way as they are commonly defined in the Finnish planning policy (see section 3.2.3). The maximum green times of the traffic-responsive signal groups were the pretimed green intervals multiplied by 1.5.

6.2 Results of the simulation study

6.2.1 Delays in intersection Basic-1

Appendix A presents the delays in a simple intersection of two one-way streets (Fig. 6.2). In traffic-responsive control the passage detectors were at distances 120 and 60 m from the stop line, supplemented by presence detectors in front of the stop lines.

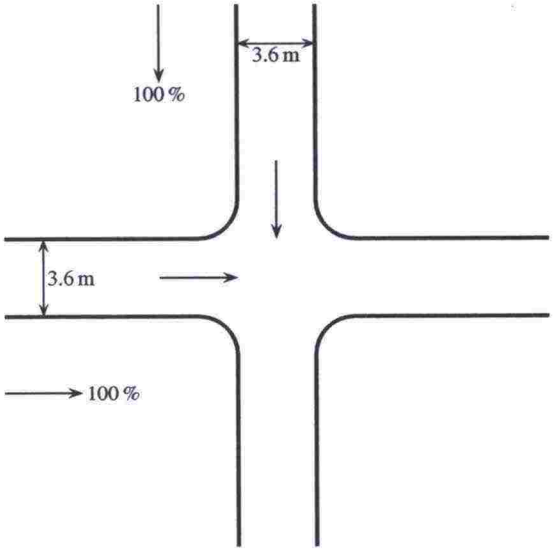


Figure 6.2: Intersection type Basic-1

At degrees of saturation below 0.5 the HCM delays for *pretimed control* are approximately equal to the Webster delays. As the degree of saturation increases above 0.5, HCM delays become higher than Webster delays.

HCM underestimates the simulated delays of pretimed control, except at degrees of saturation exceeding 0.8, where HCM delay estimates are higher than simulated delays. DanKap delays are nearly equal to HCM delays. At high degrees of saturation DanKap, however, gives lower delays, which in most cases are closer to the simulated delays. Capcal 2 delays are in most cases slightly higher than the simulated delays. At high volumes the low base saturation flow rate of Calcal 2 may be one reason for high delays.

HCM mostly overestimates *traffic-responsive control* delays at very low ($\rho < 0.4$) and high ($\rho > 0.85$) degrees of saturation. At moderate degrees of saturation HCM delays are lower than simulated delays. At high degrees of saturation the VA-delay calculation of HCM2000 appears to be more sensitive to the growth of traffic volumes than the simulation results. Under low degrees of saturation, the simulated VA-delays grow faster. The shape of the simulated delay curve is more linear, but the simulated and HCM2000 delay curves have a similar shape. To adjust the HCM2000 method to the Finnish traffic conditions, the relation between the degree of saturation and delay should be more linear.

Traffic responsive Capcal 2 delays are nearly equal to the simulated delays. At high degrees of saturation and high proportion of major flow, the major flow delays are higher and minor flow delays lower than simulated. The intersection delay is higher. This may indicate that the green split estimated by Capcal 2 for high degrees of saturation is not optimal and/or that Capcal 2 underestimated the intersection capacity.

In conclusion, all methods give reasonably good results for this intersection type. Capcal 2 gives best delay estimates, although it does overestimate delays of pretimed control at low and high degrees of saturation.

Comparison of pretimed and traffic-responsive controls indicate that the latter gives only slightly lower delays. The main advantage of traffic-responsive control is not better optimization of cycle length and green split at given conditions, nor even responsiveness to random variations, but adaptation to the systematic fluctuations of traffic flow, which

make the pretimed control not optimal for prevailing conditions.

6.2.2 Delays in intersection Basic-2

Basic-2 was an intersection of two two-lane two-way streets with no turning lanes (Fig. 6.3). In traffic-responsive control the passage detectors were at distances 120 and 60 m from the stop line, supplemented by presence detectors in front of the stop lines.

The simulations and delay calculations were carried out for two traffic conditions. In the first (Appendix B) the proportion of left turning traffic was 10 %, and in the second (Appendix C) it was 25 %. Because of the shared lanes for through and turning movements, the number of vehicles turning left has major influence on the delays of the whole approach.

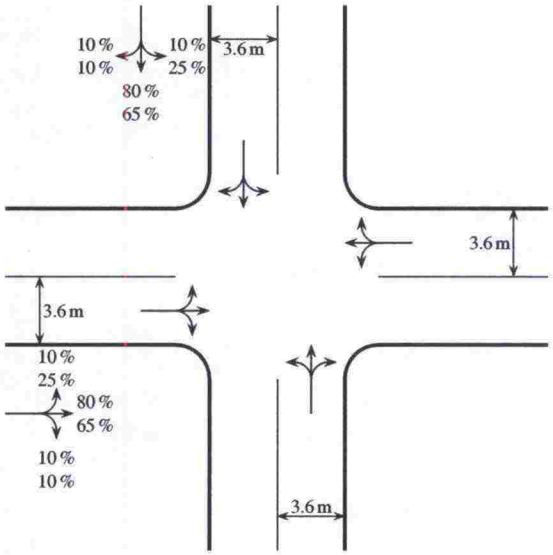


Figure 6.3: Intersection type Basic-2

With all minor/major volume combinations and degrees of saturation less than 0.75 for 10 % left-turners and less than 0.50 for 25 % left-turners the HCM2000, DanKap and Capcal 2 delays for *pretimed control* are close to the simulated delays. Webster delays are lower than the simulation results.

HCM delays are lower than simulated delays at low degrees of saturation. For all minor/major volume combinations HCM delays increase steeply after the degree of saturation exceeds 0.75. The increase is especially steep when the proportion of left-turning traffic is large. This indicates that HCM overestimates the effect of left-turning traffic. The smaller the proportion on minor flow, the closer the minor approach delays of HCM2000 are to the simulated delays. This result is probably related to the fact that the effect of left-turning traffic on the HCM-delay decreases, when the minor flow volume decreases. However, simulations indicated a higher capacity than HCM.

DanKap gives similar results to HCM2000 at low degrees of saturation in pretimed control. At high degrees of saturation DanKap delays are lower than HCM delays and closer to the simulated delays.

Capcal 2 results for pretimed control are very close to the simulated delays. At low degrees of saturation Capcal 2 overestimates the delays slightly. At high degrees of saturation the results are scattered. The modeling of permitted left turns on a shared lane appears to be more realistic in Capcal 2 than in HCM2000 and DanKap.

For *traffic-responsive control* HCM delays behave in a similar way as for pretimed control. At low degrees of saturation delays are slightly underestimated, but increase significantly above the simulated delays at high degrees of saturation.

Capcal 2 delays for traffic-responsive control have a very good agreement with the simulation results. For low percentages of left-turn traffic (Appendix B) Capcal 2 underestimates delays at high degrees of saturation ($\rho > 0.75$). The results are also somewhat inconsistent, leading to lower delays at high degrees of saturation. This inconsistency disappears as the proportion of left-turning vehicles increases (Appendix C).

The signal timings were calculated by the optimization algorithm of Capcal 2, and the program had some convergence problems. This property of Capcal 2 requires further research.

In conclusion, Capcal 2 gives the best agreement with the simulated delays. It is, however, suggested that the signal timings are either entered manually, or at least checked for consistency at high degrees of saturation, especially, if Capcal 2 reports convergence problems.

6.2.3 Delays in intersection HCM-1

Intersection type HCM-1 (Fig. 6.4) has four lanes in the major street and two lanes in the minor street, but no turning lanes. This is the intersection layout in Example Problem 1 in HCM2000. In traffic-responsive control the passage detectors were at distances 120 and 60 m from the stop line, supplemented by presence detectors in front of the stop lines.

The proportion of turning traffic was 10 percent, 5 % right and 5 % left. Appendix D presents the delays of intersection type HCM-1 according to simulations and theoretical formulas.

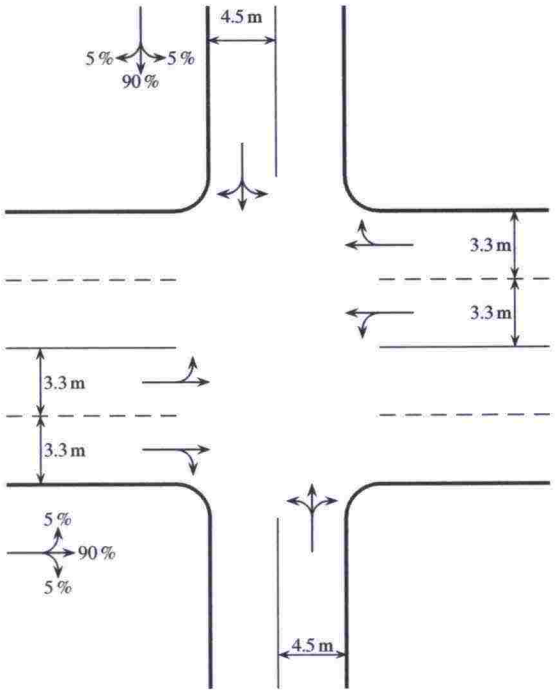


Figure 6.4: Intersection type HCM-1

In all minor/major volume combinations under *pretimed control* with degree of saturation 0.75 or lower, the HCM delay is lower than the simulated delay. When the degree of saturation is higher than 0.75 the HCM method suggests much higher delays than the simulations. Once again, the capacity according to the HCM-method is lower than in the simulations and the Webster's formula gives considerably lower delays than the other methods.

DanKap underestimates the delays on the minor approach. The delay estimates on major flow exceed the simulated delays, when both the degree of saturation and the proportion of major flow are high. As the proportion of flow on the major direction increases, DanKap underestimates the capacity in the major approach.

Capcal 2 overestimates the delays of pretimed control. At high degrees of saturation Capcal 2 delays are, however, lower than or equal to the HCM delays.

For *traffic-responsive control* the agreement between HCM2000 delays and simulated delays is better on the minor approach than on the major approach. The HCM delays estimates are closest to the simulated delays at low degrees of saturation. The major flow delay curve has a similar shape as in the Basic-2 intersection. Despite the different lane configurations in Basic-1 and Basic-2, the relation between the HCM and simulated delay curves is similar: As the degree of saturation exceeds 0.75, the HCM delay increases above the simulated delay and grows much steeper.

Capcal 2 delays for traffic responsive control are closer to the simulation results than HCM delays. Capcal 2 does, however, behave inconsistently in some situations, so that the delay increases too steeply at degrees of saturation between 0.5 and 0.75 and then decreases as the degree of saturation increases beyond 0.75. This problem is especially prominent in the minor approach delay estimates. The same phenomenon can also be observed under pretimed control, when traffic flows in the major and minor approaches are equal (Fig. D.5).

6.2.4 Delays in intersection HCM-2

Appendix E presents the delays in the intersection type HCM-2 (see Example Problem 3 of HCM2000). The intersection has two lanes plus a short exclusive left-turn lane (effective length 18 m) at each approach (Fig. 6.5). Left turns on the major (north-south) street are permitted plus protected. The protected-plus-permitted left turn suggested in the HCM example is not allowed in Finland (see Figure 2.3). In traffic-responsive control the passage detectors are at distances 120 and 60 m from the stop line, supplemented by presence detectors in front of the stop lines.

The HCM delays for *pretimed control* are only slightly lower than the simulated delays at degrees of saturation not exceeding 0.75. At higher degrees of saturation HCM underestimates the major flow delays, but gives good minor approach delay estimates, especially when the proportion of flow in the minor approach is low.

HCM underestimates the major flow capacity, when the flows in both directions are equal. As the proportion of flow on the major approach increases, HCM delays at high degrees of saturation are lower than the simulated delays. The simulation results indicate that the capacity is lower than the theoretical capacity. One reason for this may be that at high degrees of saturation and high flow rates in the major direction the left-turn lane is too short. The blocking effect reduces the capacity, as the simulation results indicate. HCM cannot properly model this blocking effect.

Under high flow rates on the minor approach, the significant effect of permitted left turns in HCM is again visible (Fig. E.5). The capacity of the HCM method is exceeded when

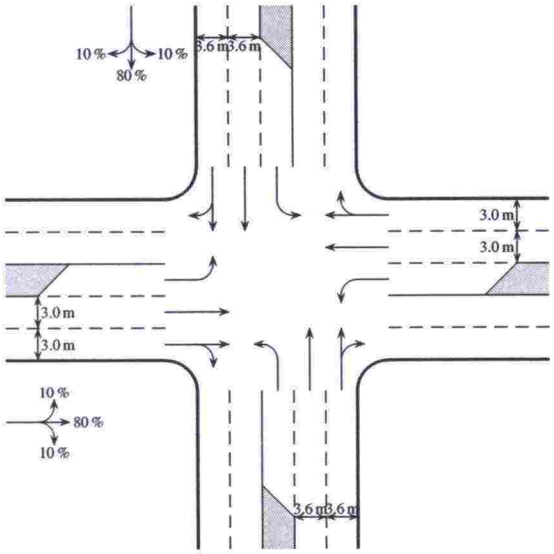


Figure 6.5: Intersection type HCM-2

the theoretical degree of saturation increases above 0.75, while the simulated control is still able to serve the traffic satisfactorily. At low proportions of minor volume this problem does not exist (see Figs. E.11 and E.17).

Capcal 2 gives slightly higher delay estimates than DanKap. Both methods are reliable, but Capcal 2 approximates the simulated delays more accurately. The increase in delay stays at a moderate level when $\rho < 0.85$, but at higher degrees of saturation ($\rho > 0.85$) the delay increases very steeply.

With equal traffic volumes on major and minor directions the HCM delay curves for *traffic-responsive control* are quite similar to the simulated delays. The distance between the curves increases with increasing degree of saturation, but the mutual order of the major flow delay curves remains unchanged. Both HCM2000 and Capcal 2 underestimate the delays of traffic responsive control. HCM2000 delay estimates are, however, closer to the simulated delays as flow rates approach capacity.

The HCM2000 delays for both pretimed and traffic-responsive control are good estimates for an approach with an exclusive left-turn lane. The protected left-turn phase (max 5 seconds) reduces the effect of left-turn movements in delay calculations and gives better correlation between simulated and HCM delays. Capcal 2 underestimates the delays of permitted-plus-protected left turns in traffic-responsive signals at high degrees of saturation ($\rho > 0.75$).

6.2.5 Delays in intersection LIVASU

Delays for intersection type LIVASU (Fig. 6.6) are presented in Appendix F. This is a typical signal controlled intersection on a Finnish multilane highway, as presented in the Finnish handbook LIVASU 95 (Kehittämiskeskus 1996). Both major street approaches have exclusive left-turn lanes with protected left-turn phases. In this example the southern leg of the intersection has only one lane for all movements, while the northern leg has an exclusive left-turn lane.

The left-turn lanes have an effective length of 51 meters, which may be considered rather short for an arterial street. Speed limit is 50 km/h, as in the other intersections above. In traffic-responsive control the passage detectors were at distances 95 and 40 m

from the stop line, supplemented by presence detectors in front of the stop lines. On the left-turn lanes the passage detectors were located at a distance of 25 m from the stop line.

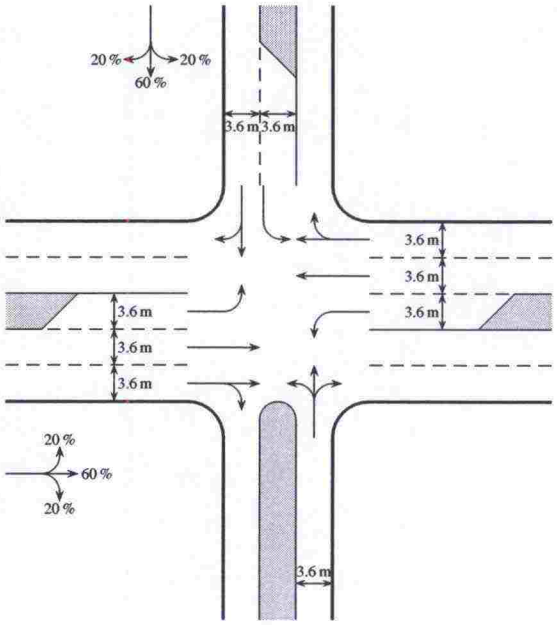


Figure 6.6: Intersection type LIVASU

The detector locations and functions are presented in Figure 3.15. Because the functions of the virtual controller of HUTSIM were more limited than the functions of a real controller, the simulated control could not reproduce all the adaptivity of the Finnish traffic-responsive control described in section 3.2.3.

Figures F.5, F.11, and F.17 for the *pretimed control* of the minor approaches indicate again that the effect of permitted left turns on the capacity in HCM are larger than those observed in the simulations. The major flow left turns are protected, and the delays in HCM and simulations are very similar.

DanKap delay curves have a similar shape as the HCM curves, but DanKap gives more accurate estimates of the simulated delays. The only significant difference is in the minor approach with a small proportion of flow (Fig. F.17). Capcal 2 gives good delay estimates for the major flow, but underestimates delays of the minor flow at high degrees of saturation.

The HCM delay curves (Figs. F.4, F.10, and F.16) for the *traffic-responsive control* of the major roadway traffic are quite similar regardless of the minor/major volume combination used, and give a reasonable estimate of the simulated delays. The delays for the minor approach (Figs. F.6, F.12, and F.18) are, however, overestimated at high degrees of saturation, as in the case of pretimed control. Capcal 2 results for traffic-responsive control are similar to the pretimed control results: good for major flow, and not so good for the minor flow, especially at high degrees of saturation.

7 CONCLUSIONS AND DISCUSSION

The object of the research was to evaluate how well HCM2000, DanKap, and Capcal 2 can estimate the delays at Finnish signalized intersections. The results of the methods were compared with the simulated delays obtained by the HUTSIM software calibrated for Finnish conditions. Five test intersections with pretimed and traffic-responsive control were analyzed at five degrees of saturation.

The degrees of saturation were calculated with applications based on the Webster method. The traffic volumes used in the analysis were defined according to these degrees of saturation. The degrees of saturation were usually significantly lower than the other methods suggested, and the highest degrees of saturation exceeded the capacity of other methods in many cases. This may be a source of some uncertainty in the results. On the other hand, the comparison of analytical methods and simulation results is still valid, because the traffic and control parameters were equal in both cases. The comparison was just made with somewhat higher degrees of saturation than was originally assumed.

The simulated delay estimates may become unreliable, as intersections become congested. Thus, the difference between the model delays and the simulated delays under heavy traffic may be smaller or larger than displayed in the figures.

The default values of each method were used for traffic and roadway conditions not specified in the test cases. The intersection types cover a variety of lane and control configurations. However, only one proportion (10 %) of heavy vehicles was used, and no alternative geometric designs (grade, lane width, curb radii, etc.) were analyzed. No pedestrian flows, parking, or bus stop operations were included in the models. The traffic conditions were assumed stationary during the analysis period.

7.1 Webster model

The Webster model (4.44) estimates the delays at pretimed intersections under stationary conditions. The model has not been designed for the analysis of traffic-responsive control or oversaturated conditions. The usefulness of the model is also questionable at degrees of saturation near unity.

The Webster model gives reasonable delays only at degrees of saturation lower than 0.75. As in other studies (Taale & van Zuylen 2001), the Webster method was found valid only at low degrees of saturation. Even then the delay estimates were lower than the simulated delays.

7.2 HCM2000

When the default saturation flows for simple lane and control configurations were replaced with those found valid in Finland, the HCM2000 delays for pretimed control gave good approximations of the simulation results. This can be seen especially in the test intersections with most separated movements (HCM-2 and LIVASU), in which the HCM delays in some cases approximated the simulations almost perfectly. The minor approach in the HCM-2 intersection (Fig. E.11) is perhaps the best example of this.

When the control method was changed from pretimed control to traffic-responsive control, the difference between the HCM delay and the simulations increased. The k-factor could not fully describe the effect of adaptive signal control. The study of Taale & van Zuylen (2001) has, however, indicated that the accuracy of HCM results increases, if the control parameter averages are collected over shorter time periods. Taale & van Zuylen also suggested that the variance in vehicle-actuated signal control

should be included in the delay estimation model.

For traffic-responsive control HCM estimated somewhat too low delays at low and especially at medium degrees of saturation. The delay estimates for pretimed control were slightly higher and, accordingly, gave better estimates for traffic-responsive delays. However, at high degrees of saturation the reduced delays due to the k-factor produced better results as compared to simulations, although the problem of too low capacity in HCM still remained.

The major problem in the HCM delay model is in the approximation of the effect of permitted left turns on shared lanes—and permitted left turns in general. The HCM2000 method underestimates the lane capacity and overestimates the delay at high degrees of saturation ($\rho > 0.75$) in almost all test intersections and minor/major flow combinations. The problem is apparently in the method of evaluating the interaction between left turning traffic and opposing traffic flows. At Finnish traffic conditions, the delays caused by the permitted left turns are lower than the HCM model suggests.—The style of driving in Finland appears to be more aggressive than in the U.S. The same phenomenon can be seen on two-lane highways, where the speed-flow curve in HCM2000 is considerably steeper than the Finnish curve (Luttinen 2001).

The HCM method does not consider the extra delay due to the blocking effect of short turning lanes. This effect is emphasized especially in the already problematic situations with high degrees of saturation and a large number of left-turning vehicles.

Similar results have been obtained also by Taale & van Zuylen (2001), who compared the HCM VA-delays to field delay measurements. Levinson & Prassas (2000) compared the capacity estimates of shared left-turn lanes with pretimed control by the 1997 HCM method, the Canadian method, the Sidra (Australian) method and a simplified method developed by Levinson. For two-lane approaches, like intersection HCM-1 (Fig. 6.4) above, the Canadian method provided higher capacities than the other methods, including HCM.

The main observations of the comparison between HCM2000 traffic-responsive delay model and the simulation results were:

1. HCM2000 produced best delay estimates at intersections with protected phases and exclusive lanes. The HCM performs best at the most simple (one-way and two-phase) and most complex (multiple and exclusive lanes, multiple phases) Finnish intersections.

2. The major problem of HCM is in the procedure for left-turn movements on a shared lane with permitted left-turn phase and high traffic volumes. HCM overestimates the disturbances caused by the opposing traffic flow to the left-turn movements.

3. A good correlation between simulations and HCM in intersections with separate lanes and phases for different movements indicated that the geometric correction terms and procedures of HCM can be applied in Finnish conditions.

4. The HCM delay estimates were too low at low and medium degrees of saturation. Because the k-factor had the greatest delay reducing effect on low degrees of saturation, the pretimed delay estimates (without the k-factor) were closer to the simulated traffic-responsive delays than the HCM delay estimates for traffic-actuated control.

The Highway Capacity and Quality of Service committee has recognized that traffic movements in a lane group with a shared lane are seldom homogeneous, and this simplifying assumption in HCM2000 has an adverse effect on the accuracy and reliability of the results. The committee also aims to improve the consistency of the "end points" of various models. For example, one would expect that a permitted left turn with no opposing traffic should have the same capacity as a protected left turn. Also, in some cases adding a left-turn lane in HCM calculations appears to make matters worse, not better. The committee has formulated these issues as research problem statements.

Control delay is used as a service measure in HCM2000. There is an ongoing discussion in the Highway Capacity and Quality of Service committee on the role of the level-of-service concept. One line of discussion concerns the question, whether the quality of traffic conditions should be described in terms of six levels (A–F) of service. Another line of discussion concerns the service measures. Some studies of users' perception have indicated that control delay and six levels of service may not be a sufficient way of describing the quality of service (Pécheux et al. 2000). This subject requires further research and discussion.

7.3 DanKap

The results of DanKap are similar to the results of HCM2000, as can be expected. The delay estimates of DanKap are, however, closer to the simulation results. Also, the capacity estimates of DanKap appear to be in a better agreement with the simulation results. Evidently, the DanKap model for permitted left turns has been adjusted for Nordic conditions. For pretimed control in the conditions analyzed, DanKap results are in better agreement with the simulated delays than the results of HCM2000. DanKap does not have any adjustment for traffic-responsive control.

7.4 Capcal 2

The interaction delays (stop times) of Capcal 2 were larger than the simulated control delays at pretimed intersections under low degrees of saturation. The minimum cycle length restriction in Capcal 2 gives only a partial explanation for this deviation. The error is largest in intersections Basic-1 and HCM-1. Basic-1 has no turning traffic. Intersection HCM-1 has five percent left turning vehicles as well as right turning vehicles. In other intersections the proportion of turning traffic is higher. When the proportion of (left) turning traffic is high, Capcal 2 gives better results.

The basic saturation flow of Capcal 2 for through traffic under ideal conditions is 1,850 veh/h/lane. In Finland the saturation flow estimate under ideal conditions is 1,940 veh/h/lane. This difference can be seen as higher delays in Capcal 2 at high degrees of saturation, especially in intersection HCM-1. At the minor approach in intersection LIVASU, Capcal 2 gives very low delays at high degrees of saturation. In a smaller scale the same phenomenon can also be seen in the minor approach of intersection Basic-2 with 25 % left turners. Both of theses cases have high percentages of left turners and a shared lane for all movements. (The northern leg of intersection LIVASU has a short exclusive left-turn lane.) Consequently, Capcal 2 seems to underestimate the capacity of through traffic, and overestimate the capacity of permitted left turns. For many practical situations these errors cancel each other resulting in good delay estimates.

Capcal 2 underestimates the delays of traffic-responsive control at high degrees of saturation in intersection HCM-2 and in the minor approach of intersection LIVASU. These cases have either permitted or permitted-plus-protected control for exclusive left-turn

lanes. These cases also indicate that Capcal 2 overestimates the capacity of permitted left turning movements from exclusive turning lanes.

Capcal 2 had some convergence problems when estimating the signal timing under traffic-responsive control. Consequently, the results had some inconsistencies—delays decreased as the degree of saturation increased. It is necessary to enter the timing information manually, or at least check that the Capcal 2 timings are reasonable. This issue demands further research.

Capcal 2 uses detailed information of intersection geometry to adjust the delay estimates. Because of the limitations in HUTSIM the effects of roadway and intersection geometry could not be analyzed. The adjustments for grades and heavy vehicles, however, appear to be compatible with the Finnish measurements.

7.5 Recommendations

For pretimed control DanKap gives better delays estimates than HCM2000. DanKap does not, however, estimate the effect of traffic-responsive control. Because the difference in delay between an optimal pretimed control and a traffic-responsive control is small, and HCM and DanKap underestimate the delays at low and medium degrees of saturation, the DanKap delay estimates can be used also in the analysis of traffic-responsive systems. This analysis cannot, however, give any estimate of the effect of traffic-responsiveness for stationary conditions. The major benefits of traffic-responsive control are, however, obtained under non-stationary conditions.

Overall, the interaction delay estimates of Capcal 2 were closest to the simulated control delays. Of the three methods analyzed, Capcal 2 can be suggested as the best tool for the analysis of Finnish signalized intersections, both pretimed and traffic responsive. Capcal 2 does, however, have some convergence problems in the analysis of traffic-responsive control, and it is suggested that the cycle length and green splits are either entered manually, or at least checked for consistency.

The major advantage of traffic-responsive control is its adaptiveness to changes in traffic demand. Under stationary conditions, with only random variation, the difference in control delay between traffic-responsive control and optimized pretimed control is small. In order to obtain an estimate of the advantage of traffic-responsive control over pretimed control, the analysis should be performed for different traffic conditions, for which the actual (on-field) pretimed control is not always optimal.

"Level of service" should be preferred over "operational quality" as a quality of service indicator. The average control delay should be used as a service measure, and the LOS criteria should be those of HCM2000 (Tables 4.1 and 4.3).

In operational analysis the saturation flow rates in Figure 5.3 should be used. These values are based on measurements in the Helsinki metropolitan area and supplementary simulations. For more reliable results data collection should also be made in other parts of the country.

The current measurements are not adequate for the estimation of the effect of heavy vehicles. For the time being, the Capcal 2 model (5.13) for grade and heavy vehicle adjustment is suggested.

No guidelines are presented for the capacity analysis of "free right" turning movements. A procedure similar to the analysis of permitted left turns should be used, where the major flow is divided into periods of saturated flow, unsaturated flow and no flow. During a conflicting saturated-flow period no right turn movements are possible on the "free right" lane. During an unsaturated period (green time after queue discharge) in the

major direction the vehicles on a "free right" lane merge observing a gap-acceptance process. During a no-flow period in the major direction the vehicles on a "free right" lane can depart unaffected by conflicting flows.

There is no definite guideline for the capacity estimation of Finnish signalized intersections. The current manual (Kehittämiskeskus 1996) presents both old and new saturation flow rates, but prefers the old values. It is obvious that this method underestimates the capacity of signalized facilities, and may suggest unnecessary investments. The application of adjustment factors for the effects of turning movements, opposing traffic and pedestrian conflicts would make the method more flexible and methodologically similar to the current international methods. In addition, the service measures and LOS criteria should be more clearly defined. There is an obvious need for new guidelines.

REFERENCES

AKÇELIK, R. 1980. *Time-Dependent Expressions for Delay, Stop Rate and Queue Length at Traffic Signals*. (Internal Report AIR–367). Vermont South, Victoria: Australian Road Research Board.

AKÇELIK, R. 1981. *Traffic Signals: Capacity and Timing Analysis*. (Research Report ARR No 123). Vermont South, Victoria: Australian Road Research Board.

AKÇELIK, R. 1988. The Highway Capacity Manual delay formula for signalized intersections. *ITE Journal*, vol. 58, no. 3, p. 23–27.

ALLSOP, R. 1972. Delay at a fixed time traffic signal—I: Theoretical analysis. *Transportation Science*, vol. 6, no. 3, p. 260–285.

ALMOND, J. (ed.) 1965. *Proceedings of The Second International Symposium on The Theory of Road Traffic Flow*. Paris: The Organisation for Economic Co-operation and Development.

AUSTROADS 1988. *Roadway Capacity*. (Guide to Traffic Engineering Practice—Part 2). Sydney.

BÅNG, K.-L. 1976. Optimal control of isolated signals. *Traffic Engineering and Control*, vol. 17, no. 7, p. 288–292.

BARAS, J. S., DORSEY, A. J. & LEVINE, W. S. 1979. Estimation of traffic platoon structure from headway statistics. *IEEE Transactions on Automatic Control*, vol. 24, no. 4, p. 553–559.

BARAS, J. S. & LEVINE, W. S. 1979. Algorithms for the computer control of urban traffic. In Homburger & Steinman (1979), p. 74–91.

BARAS, J. S., LEVINE, W. S., DORSEY, A. J. & LIN, T. L. 1979. Traffic estimation based on point process models. In Levine, Lieberman & Fearnsides (1979), p. 266–285.

BARAS, J. S., LEVINE, W. S. & LIN, T. L. 1979. Discrete-time point processes in urban traffic queue estimation. *IEEE Transactions on Automatic Control.*, vol. 24, no. 1, p. 12–27.

BECKMANN, M., MCGUIRE, C. B. & WINSTEN, C. B. 1956. *Studies in the Economics of Transportation*. New Haven: Yale University Press.

BELLMAN, R. 1957. *Dynamic Programming*. Princeton, NJ: Princeton University Press.

BELL, M. G. H. 1990. A probabilistic approach to the optimisation of traffic signal settings in discrete time. In Koshi (1990), p. 619–632.

BETRÓ, B., SCHOEN, F. & SPERANZA, M. G. 1987. A stochastic environment for the adaptive control of single intersections. In Gartner & Wilson (1987).

BRIGGS, T. 1977. Time headways on crossing the stop-line after queueing at traffic lights. *Traffic Engineering and Control*, vol. 18, no. 5, p. 264–265.

BRILON, W., GROSSMANN, M. & BLANKE, H. 1994. *Verfahren für die Berechnung der Leistungsfähigkeit und Qualität des Verkehrsablaufes auf Straßen*. (Forschung Straßenbau und Straßenverkehrstechnik, Heft 669). Bonn: Bundesministerium für Verkehr.

BRILON, W. & WU, N. 1990. Delays at fixed-time traffic signals under time-dependent traffic conditions. *Traffic Engineering and Control*, vol. 31, no. 12, p. 623–631.

BURROW, I. 1989. A note on traffic delay formulas. *ITE Journal*, vol. 59, no. 10, p. 29–32.

CARDEN, P. & MCDONALD, M. 1985. The application of SCOOT control to an isolated intersection. *Traffic Engineering and Control*, vol. 26, no. 6, p. 304–310.

CATLING, I. 1977. A time-dependent approach to junction delays. *Traffic Engineering and Control*, vol. 18, no. 11, p. 520–523, 526.

CHEN, H., COHEN, S. L., GARTNER, N. H. & LIU, C. C. 1987. Simulation study of OPAC: A demand-responsive strategy for traffic signal control. In Gartner & Wilson (1987), p. 233–249.

CHODUR, J. & TRACZ, M. 1984. Study and numerical modelling of non-stationary traffic flow demands at signalized intersections. In J. VOLMULLER & R. HAMERSLAG (ed.), *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory*. Utrecht: VNU Science Press, p. 133–154.

CLAYTON, A. J. H. 1941. Road traffic calculations. *Journal of the Institution of Civil Engineers*, vol. 16, no. 7, p. 247–284. With discussion and correspondence in no. 8, pp. 588–594.

COURAGE, K. G. & PAPAPANOU, P. 1977. Estimation of delay at traffic-actuated signals. In *Transportation Research Record 1365*. Washington, D.C.: Transportation Research Board. p. 17–21.

CRONJÉ, W. B. 1983. Derivation of equations for queue length, stops, and delay for fixed-time traffic signals. In *Traffic Flow, Capacity, and Measurements*. (Transportation Research Record 905). Washington, D.C.: Transportation Research Board. p. 93–95.

DARROCH, J. N. 1964. On the traffic-light queue. *Annals of Mathematical Statistics*, vol. 35, p. 380–388.

DARROCH, J. N., NEWELL, G. F. & MORRIS, R. W. J. 1964. Queues for a vehicle-actuated traffic light. *Operations Research*, vol. 12, no. 6, p. 882–895.

DE SMIT, J. H. A. 1971. The transient behaviour of the queue at a fixed cycle traffic light. *Transportation Research*, vol. 5, no. 1, p. 1–14.

DUNNE, M. C. 1967. Traffic delay at a signalized intersection with binomial arrivals. *Transportation Science*, vol. 1, no. 1, p. 24–31.

DUNNE, M. C. & POTTS, R. B. 1964. Algorithm for traffic control. *Operations Research*, vol. 12, no. 6, p. 870–881.

DUNNE, M. C. & POTTS, R. B. 1967. Analysis of a computer control of an isolated intersection. In Edie, Herman & Rothery (1967), p. 258–266.

EDIE, L. C., HERMAN, R. & ROTHERY, R. (ed.) 1967. *Vehicular Traffic Science*. (Proceedings of the Third International Symposium on the Theory of Traffic Flow, New York, June 1965). New York: Elsevier.

ELORANTA, T. 1998. Liikennetieto-ohjauksen periaatteet [Principles of traffic-responsive control]. In M. PURSULA, T. LUTTINEN & J. OJALA (ed.), *Liikenteen palvelutaso ja välityskyky* [Level of service and capacity of vehicular traffic]. (Julkaisu 95). Otaniemi: Teknillinen korkeakoulu. Chapter V.

ENGLEBRECHT, R. J., FAMBRO, D. B., ROUPHAIL, N. M. & BARKAWI, A. A. 1997. Validation of generalized delay model for oversaturated conditions. In *Highway Capacity Issues and Analysis*. (Transportation Research Record 1572). Washington, D.C.: Transportation Research Board. p. 122–130.

EVANS, H. K. (ed.) 1950. *Traffic Engineering Handbook*. Second ed. New Haven, Connecticut: Institute of Traffic Engineers.

FAMBRO, D. B. & ROUPHAIL, N. M. 1997. Generalized delay model for signalized intersections and arterial streets. In *Highway Capacity Issues and Analysis*. (Transportation Research Record 1572). Washington, D.C.: Transportation Research Board. p. 112–121.

GARTNER, N. H. 1983. OPAC: A demand-responsive strategy for traffic signal control. In *Transportation Research Record 906*. Washington, D.C.: Transportation Research Board. p. 75–81.

GARTNER, N. H. & WILSON, N. H. M. (ed.) 1987. *Transportation and Traffic Theory*. (Proceedings of the Tenth International Symposium on Transportation and Traffic Theory). New York: Elsevier.

GARWOOD, F. 1940. An application of the theory of probability to the operation of vehicular-controlled traffic signals. *Journal of the Royal Statistical Society*, vol. 7, no. 1, p. 65–77.

GAZIS, D. C. 1964. Optimum control of a system of oversaturated intersections. *Operations Research*, vol. 12, p. 815–831.

GAZIS, D. C. & POTTS, R. B. 1965. The over-saturated intersection. In Almond (1965), p. 221–237.

GORDON, R. 1969. A technique for control of traffic at critical intersections. *Transportation Science*, vol. 4, no. 3, p. 279–288.

GRAFTON, R. B. & NEWELL, G. F. 1967. Optimal policies for the control of an undersaturated intersection. In Edie et al. (1967), p. 239–257.

HAIGHT, F. A. 1959. Overflow at a traffic light. *Biometrika*, vol. 46, p. 420–424.

HAIGHT, F. A. 1963. *Mathematical Theories of Traffic Flow*. New York: Academic Press.

HAMMOND, H. F. & SORENSON, L. J. (ed.) 1941. *Traffic Engineering Handbook*. New York: Institute of Traffic Engineers and National Conservation Bureau.

HEYDECKER, B. 1987. Uncertainty and variability in traffic signal calculations. *Transportation Research*, vol. 21B, no. 1, p. 79–85.

HEYDECKER, B. G. 1990. Continuous-time formulation for traffic-responsive signal control. In Koshi (1990), p. 599–618.

HOMBURGER, W. S. & STEINMAN, L. (ed.) 1979. *Proceedings of the International Symposium on Traffic Control Systems*, Vol. 2A. University of California at Berkeley, Institute of Transportation Studies and the U.S. Department of Transport, Berkeley, Ca.

HUNT, P. B., ROBERTSON, D. I., BRETHERTON, R. D. & ROYLE, M. C. 1982*a*. The SCOOT on-line traffic signal optimization technique. In *International Conference on Road Traffic Signalling* (The Institution of Electrical Engineers 1982), p. 59–62.

HUNT, P. B., ROBERTSON, D. I., BRETHERTON, R. D. & ROYLE, M. C. 1982*b*. The SCOOT on-line traffic signal optimization technique. *Traffic Engineering and Control*, vol. 23, no. 4, p. 190–192.

HUNT, P. B., ROBERTSON, D. I., BRETHERTON, R. D. & WINTON, R. I. 1981. *SCOOT—a traffic responsive method of coordinating signals*. (TRRL Laboratory Report 1014). Crowthorne, Berkshire: Transport and Road Research Laboratory.

HURDLE, V. F. 1984. Signalized intersection delay model—A primer for the uninitiated. In *Traffic Capacity and Characteristics*. (Transportation Research Record 971). Washington, D.C.: Transportation Research Board. p. 96–105.

HUTCHINSON, T. P. 1972. Delay at a fixed time traffic signal—II: Numerical comparisons of some theoretical expressions. *Transportation Science*, vol. 6, no. 3, p. 286–305.

JOVANIS, P. P. & GREGOR, J. A. 1986. Coordination of actuated arterial traffic signal systems. *Journal of Transportation Engineering*, vol. 112, no. 4, p. 416–432.

KEHITTÄMISKESKUS 1996. *LIVASU 95—Liikennevalot* [Traffic signal handbook]. Helsinki: Tielaitos.

KELL, J. H. & FULLERTON, I. J. 1982. *Manual of Traffic Signal Design*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

KENDALL, D. G. 1951. Some problems in the theory of queues. *Journal of the Royal Statistical Society*, vol. 13, no. 2, p. 151–185. With discussion.

KHATIB, Z. K. & KYTE, M. 2001. Uncertainty in projecting the level of service of signalized and unsignalized intersections. In *80th Annual Meeting Preprint CD-ROM*. Washington, D.C.: Transportation Research Board.

KIMBER, R. M. & HOLLIS, E. M. 1978. Peak-period traffic delays at road junctions and other bottlenecks. *Traffic Engineering and Control*, vol. 19, no. 10, p. 442–446.

KIMBER, R. M. & HOLLIS, E. M. 1979. *Traffic queues and delays at road junctions*. (TRRL Laboratory Report 909). Crowthorne, Berkshire: Transport and Road Research Laboratory.

KIMBER, R. M., MARLOW, M. & HOLLIS, E. M. 1977. Flow/delay relationships for major/minor priority junctions. *Traffic Engineering and Control*, vol. 18, no. 11, p. 516–519.

KING, G. F. & WILKINSON, M. 1976. Relationship of signal design to discharge headway, approach capacity, and delay. In *Capacity and Measurement of Effectiveness*. (Transportation Research Record 615). Washington, D.C.: Transportation Research Board. p. 37–44.

KITTELSON, W. K. & ROESS, R. P. 2001. Highway capacity analysis after Highway Capacity Manual 2000. In *Traffic Flow Theory and Highway Capacity 2001*. (Transportation Research Record 1776). Washington, D.C.: Transportation Research Board. p. 10–16.

KIVELÄ, M. & PURSULA, M. 1982. *Valo-ohjauksisen tasoliittymän välityskyky ja viivytyslaskelmat—Ruotsalainen menetelmä* [Capacity and delay calculations for a signalized intersection—The Swedish method]. (Liikennetekniikka, Opetusmoniste 1). Otaniemi: Helsingin teknillinen korkeakoulu.

KOSHI, M. 1979. A new method of traffic-responsive control of traffic signals. In Homburger & Steinman (1979), p. 59–73.

KOSHI, M. (ed.) 1990. *Transportation and Traffic Theory*. (Proceedings of the Eleventh International Symposium on Transportation and Traffic Theory). New York: Elsevier.

KOSONEN, I. 1996. *HUTSIM—Simulation Tool for Traffic Signal Control Planning*. (Transportation Engineering Publication 89). Otaniemi: Helsinki University of Technology.

KOSONEN, I. 1999. *HUTSIM—Urban Traffic Simulation and Control Model: Principles and Applications*. (Transportation Engineering Publication 100). Espoo: Helsinki University of Technology.

KRONBORG, P. 1992. *MOVA and LHOVRA Traffic Signal Control for Isolated Intersections*. (TFK Report 1992:4E). Stockholm: Transport Research Institute.

KRONBORG, P., DAVIDSSON, F. & EDHOLM, J. 1997. *SOS—Self Optimising Signal Control*. (TFK Report 1997:2E). Stockholm: Transport Research Institute.

LEHOCZKY, J. P. 1972. Traffic intersection control and zero-switch queues under conditions of markov chain dependence input. *Journal of Applied Probability*, vol. 9, p. 382–395.

LEUTZBACH, W. & KÖHLER, U. 1974. Definitions and relationships for three different time intervals for delayed vehicles. In D. J. BUCKLEY (ed.), *Transportation and Traffic Theory*. (Proceedings of the Sixth International Symposium on Transportation and Traffic Theory). New York: Elsevier, p. 87–103.

LEVINE, W. S., LIEBERMAN, E. & FEARNSIDES, J. J. (ed.) 1979. *Research Directions in Computer Control of Urban Traffic Systems*. (Proceedings Engineering Foundation Conference). American Society of Civil Engineers, New York.

LEVINSON, H. S. & PRASSAS, E. S. 2000. The capacity of shared left turn lanes—A comparative analysis. In *79th Annual Meeting Preprint CD-ROM*. Washington, D.C.: Transportation Research Board.

LIIKENNE- JA VIESTINTÄMINISTERIÖ 2000. *Tieliikenteen liikennevalot—Työryhmän tarkistetut ehdotukset* [Traffic signals in road traffic—Revised proposals of a working group]. (Mietintöjä ja muistioita B 26/2000). Helsinki.

LI, J., ROUPHAIL, N. M. & AKÇELIK, R. 1994. Overflow delay estimation for a simple intersection with fully actuated signal control. In *Part 2: Traffic Flow and Capacity*. (Transportation Research Record 1457). Washington, D.C.: Transportation Research Board. p. 73–81.

LIN, F.-B. 1983. Delay models of traffic-actuated signal controls. In *Traffic Flow, Capacity, and Measurements*. (Transportation Research Record 905). Washington, D.C.: Transportation Research Board. p. 33–38.

LITTLE, J. D. C. 1961. A proof of the queuing formula $L = \lambda W$. *Operations Research*, vol. 9, no. 3, p. 383–387.

LITTLE, JR., J. G. 1971. Queuing of side-street traffic at a priority type vehicle-actuated signal. *Transportation Research*, vol. 5, p. 295–300.

LOWRIE, P. R. 1982. The Sydney co-ordinated adaptive traffic system—principles, methodology, algorithms. In *International Conference on Road Traffic Signalling* (The Institution of Electrical Engineers 1982), p. 67–70.

LUK, J. Y. K. 1984. Two traffic-responsive Area Traffic Control methods: SCAT and SCOOT. *Traffic Engineering and Control*, vol. 25, no. 1, p. 14–18, 22.

LUTTINEN, R. T. 1994. Liikennetieto-ohjaus keskusta-alueilla [Traffic-responsive control in urban areas]. In *Tie ja Liikenne '94*. Tampere: Suomen Tieyhdistys.

LUTTINEN, R. T. 1996. *Statistical Analysis of Vehicle Time Headways*. (Transportation Engineering Publication 87). Otaniemi: Helsinki University of Technology.

LUTTINEN, R. T. 2001. *Capacity and Level of Service on Finnish Two-Lane Highways*. (Finnra reports 18/2001). Helsinki: Finnish Road Administration.

LUTTINEN, T. 1988. *Alueellisen valo-ohjauksen kehittäminen FTC-12000 ympäristössä* [Development of area traffic control in FTC-12000 environment]. Lahti: Insinööritoimisto TL-Suunnittelu Oy.

LYLY, S. (ed.) 1988. *Liikenne ja väylät* [Transportation Planning anf Facilities]. Helsinki: Suomen Rakennusinsinöörien Liitto. In two volumes: RIL 165-I and RIL 165-II.

MÄENPÄÄ, M. 2000. Sumea valo-ohjaus ja joukkoliikenne-etuudet [Fuzzy signal control and public transport priorities]. Master's thesis. Espoo: Helsinki University of Technology, Department of Civil and Environmental Engineering.

MÄKELÄ, K. 1997. Liikennevalo-ohjauksen palvelutasoluokitus [Level of service criteria for signalized intersections]. Master's thesis. Tampere: Tampere University of Technology.

MAY, A. D. & KELLER, H. E. M. 1967. A deterministic queueing model. *Transportation Research*, vol. 1, no. 2, p. 117–128.

MCNEIL, D. R. 1968. A solution to the fixed-cycle traffic light problem for compound Poisson arrivals. *Journal of Applied Probability*, vol. 5, p. 624–635.

MICHALOPOULOS, P. G., PAPAPANOU, B. & BINSEEL, E. B. 1978. Performance evaluation of traffic actuated signals. *Transportation Engineering Journal of ASCE*, vol. 104, no. TE5, p. 621–636.

MICHALOPOULOS, P. G. & STEPHANOPOULOS, G. 1979. An algorithm for real-time control of critical intersections. *Traffic Engineering and Control*, vol. 20, no. 1, p. 9–15.

MILLER, A. J. 1963. Settings for fixed-cycle traffic signals. *Operational Research Quarterly*, vol. 14, no. 4, p. 373–386.

MILLER, A. J. 1965. A computer control system for traffic networks. In Almond (1965), p. 200–220.

MORRIS, R. W. J. & PAK-POY, P. G. 1967. Intersection control by vehicle actuated signals. *Traffic Engineering and Control*, vol. 8, no. 10, p. 288–293.

NEUBURGER, H. 1971. The economics of heavily congested roads. *Transportation Research*, vol. 5, no. 4, p. 283–293.

NEWELL, G. F. 1960. Queues for a fixed-cycle traffic light. *Annals of Mathematical Statistics*, vol. 31, p. 589–597.

NEWELL, G. F. 1965. Approximation methods for queues with application to the fixed-cycle traffic light. *SIAM Review*, vol. 7, no. 2, p. 223–240.

NEWELL, G. F. 1969. Properties of vehicle-actuated signals: I. One-way streets. *Transportation Science*, vol. 3, no. 1, p. 30–52.

NEWELL, G. F. 1982. *Applications of Queueing Theory*. Second ed. London: Chapman and Hall.

NEWELL, G. F. 1989. *Theory of Highway Traffic Signals*. (Course Notes UCB-ITS-CN-89-1). Berkeley, Ca.: Institute of Transportation Studies, University of California at Berkeley.

NEWELL, G. F. 1990. Stochastic delays on signalized arterial highways. In Koshi (1990), p. 589–598.

NEWELL, G. F. & OSUNA, E. E. 1969. Properties of vehicle-actuated signals: II. Two-way streets. *Transportation Science*, vol. 3, no. 2, p. 99–125.

NIITTYMÄKI, J. 1993. Calibration of HUTSIM traffic signal simulator. In *Simulering och Trafiksignalstyrning—Seminariet* [Simulation and traffic signal control—Seminar]. TFK. Stockholm.

NIITTYMÄKI, J. 1998. *Isolated Traffic Signals — Vehicle Dynamics and Fuzzy Control*. (Transportation Engineering Publication 94). Otaniemi: Helsinki University of Technology.

NIITTYMÄKI, J. 2002. *Fuzzy Traffic Signal Control—Principles and Applications*. (Transportation Engineering Publication 103). Espoo: Helsinki University of Technology.

NIITTYMÄKI, J. & PURSULA, M. 1994. *Valo-ohjattujen liittymien simulointi ja ajodynamiikka* [Simulation and vehicle dynamics of signalized intersections]. (Julkaisu 81). Otaniemi: Teknillinen korkeakoulu, Liikennetekniikka.

NIITTYMÄKI, J. & PURSULA, M. 1997. Saturation flows at signal-group-controlled traffic signals. In *Highway Capacity Issues and Analysis*. (Transportation Research Record 1572). Washington, D.C.: Transportation Research Board. p. 24–32.

OHNO, K. 1978. Computational algorithm for a fixed cycle traffic signal and new approximate expressions for average delay. *Transportation Science*, vol. 12, no. 1, p. 29–47.

OLSZEWSKI, P. S. 1990. Modelling of queue probability distribution at traffic signals. In Koshi (1990), p. 569–588.

ORCUTT, JR., F. L. 1993. *The Traffic Signal Book*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

PÉCHEUX, K. K., PIETRUCHA, M. T. & JOVANIS, P. P. 2000. User perception of level of service at signalized intersections: Methodological issues. In W. BRILON (ed.), *Fourth International Symposium on Highway Capacity Proceedings*. (Transportation Research Circular E-C018). Washington, D.C.: Transportation Research Board, p. 322–335.

PEIRCE, J. R. & VINCENT, R. A. 1989. Repeated-stage signal control using MOVA. *Traffic Engineering and Control*, vol. 30, no. 1, p. 6–9.

PETERSON, A., BERGH, T. & STEEN, K. 1986. LHOVRA—a new traffic signal control strategy for isolated junctions. *Traffic Engineering and Control*, vol. 27, no. 7/8, p. 388–389.

PIGNATARO, L. J. 1973. *Traffic Engineering: Theory and Practice*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

POHJOISMAIDEN TIETEKNILLINEN LIITTO 1978. *Liikenteen valo-ohjauksen suunnittelu LIVASU -78* [Design of traffic signal control]. (Raportti 1:1978). N.p.

POTTS, R. B. 1967. Traffic delay at a signalized intersection with binomial arrivals. *Transportation Science*, vol. 1, no. 2, p. 126–128.

PURSULA, M. & NIITTYMÄKI, J. 1996. Evaluation of traffic signal control with simulation—a comparison of the Pappis-Mamdani fuzzy control vs. vehicle actuation with the extension principle. In *4th Meeting of the EURO Working Group on Transportation*. University of Newcastle upon Tyne.

ROBERTSON, D. I. 1969. *TRANSYT: A Traffic Network Study Tool*. (RRL Report LR 253). Crowthorne, Berkshire: Road Research Laboratory.

ROSDOLSKY, H. G. 1973. A method for adaptive traffic control. *Transportation Research*, vol. 7, no. 1, p. 1–15.

ROUPHAIL, N. M. & AKÇELIK, R. 1992. Oversaturation delay estimates with consideration of peaking. In *Highway Capacity and Traffic Flow*. (Transportation Research Record 1365). Washington, D.C.: Transportation Research Board. p. 71–81.

ROUPHAIL, N. M., ANWAR, M., FAMBRO, D. B., SLOUP, P. & PEREZ, C. E. 1997. Validation of generalized delay model for vehicle-actuated traffic signals. In *Highway Capacity Issues and Analysis*. (Transportation Research Record 1572). Washington, D.C.: Transportation Research Board. p. 105–111.

ROUPHAIL, N. M., TARKO, A. & LI, J. 1997. Traffic flow at signalized intersections. In N. H. GARTNER & C. J. MESSER (ed.), *Monograph on Traffic Flow Theory*. Np.: Federal Highway Administration. Chapter 9. http://www.tfhrc.gov/its/tft/tft.htm, 08/03/2001.

SANE, K. J. 1986. *Liikennetieto-ohjattujen liikennevalojen toimintaperiaatteita* [Operation of traffic-responsive signals]. Helsinki: Tie- ja vesirakennushallitus, Tiensuunnittelutoimisto.

SHAWALY, E. A. A., ASHWORTH, R. & LAURENCE, C. J. D. 1988. A comparison of observed, estimated and simulated queue lengths and delays at oversaturated signalized junctions. *Traffic Engineering and Control*, vol. 29, no. 12, p. 637–643.

SIMS, A. G. 1979. The Sydney co-ordinated adaptive traffic system. In Levine et al. (1979), p. 12–27.

SMITH, M., CLEGG, J. & YARROW, R. 2001. Modeling traffic signal control. In K. J. BUTTON & D. A. HENSHER (ed.), *Handbook of Transport Systems and Traffic Control*. Vol. 3 of *Handbooks in Transport*. Amsterdam: Pergamon. p. 503–525.

STATENS VÄGVERK 1977. *Beräkning av kapacitet, kölängd, fördröjning i vägtrafikanläggningar* [Calculation of capacity, queue length, delay in road traffic facilities]. Stockholm. TV 131.

SWEDISH NATIONAL ROAD ADMINISTRATION 1995. *CAPCAL 2: Model description of Intersection with signalcontrol*. Borlänge. Publication 1995:008E.

TAALE, H. & VAN ZUYLEN, H. J. 2001. Testing the HCM 1997 delay function for Dutch signal controlled intersections. In *80th Annual Meeting Preprint CD-ROM*. Washington, D.C.: Transportation Research Board.

TEPLY, S. 1981. Saturation flow at signalized intersections through a magnifying glass. In V. F. HURDLE, E. HAUER & G. N. STEUART (ed.), *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory*. Toronto: University of Toronto Press, p. 588–622.

TEPLY, S. 1989. Accuracy of delay surveys at signalized intersections. In *Highway Capacity, Flow Measurement, and Theory*. (Transportation Research Record 1225). Washington, D.C.: Transportation Research Board. p. 24–32.

TEPLY, S., ALLINGHAM, D. I., RICHARDSON, D. B. & STEPHENSON, B. W. 1995. *Canadian Capacity Guide for Signalized Intersections*. Np.: Institute of Transportation Engineers, District 7 – Canada.

TEPLY, S. & JONES, A. M. 1991. Saturation flow: Do we speak the same language? In *Freeway Operations, Highway Capacity, and Traffic Flow*. (Transportation Research Record 1320). Washington, D.C.: Transportation Research Board. p. 144–153.

TERVALA, J. & APPEL, K. 1987. New signal control strategies on high-speed roads in Finland. *Traffic Engineering and Control*, vol. 28, no. 3, p. 114–118.

THE INSTITUTION OF ELECTRICAL ENGINEERS 1982. *International Conference on Road Traffic Signalling*. London.

THE UNIVERSITY OF FLORIDA TRANSPORTATION RESEARCH CENTER 1996. *Capacity Analysis of Traffic-Actuated Intersections*. (NCHRP Web Document 10). Np.: National Cooperative Highway Research Program, TRB, National Research Council. NCHRP Project 3-48 Final Report.

TIENSUUNNITTELUTOIMISTO 1987. *CAPCAL-tietokoneohjelman käyttöohje 2* [CAPCAL user's manual 2]. Helsinki: Tie- ja vesirakennushallitus. TVH 723860.

TRANSPORTATION RESEARCH BOARD 1985. *Highway Capacity Manual*. (Special Report 209). Washington, D.C.: National Research Council.

TRANSPORTATION RESEARCH BOARD 1994. *Highway Capacity Manual.* (Special Report 209). Third ed. Washington, D.C.: National Research Council. Updated October 1994.

TRANSPORTATION RESEARCH BOARD 1998. *Highway Capacity Manual.* (Special Report 209). Third ed. Washington, D.C.: National Research Council. Updated December 1997.

TRANSPORTATION RESEARCH BOARD 2000. *Highway Capacity Manual.* Washington, D.C.: National Research Council. HCM2000, metric units.

VÄGVERKET 1983. *Signalreglering med LHOVRA-teknik—Projekteringshandbok* [Signal control with the LHOVRA method—Project handbook]. Borlänge. TU 155.

VÄGVERKET 1995. *CAPCAL 2: Manual för programkörning* [CAPCAL 2: User manual]. Borlänge. Publication 1995:006.

VAN ZIJVERDEN, J. D. & KWAKERNAAK, H. 1969. A new approach to traffic-actuated computer control of intersections. In W. LEUTZBACH & P. BARON (ed.), *Beiträge zur Theorie des Verkehrsflusses.* (Strassenbau und Strassenverkehrstechnik 86). Bundesminister für Verkehr. Bonn: Bundesminister für Verkehr, p. 113–117.

VEJDIREKTORATET 1999*a*. *Kapacitet og serviceniveau: Baggrund og dokumentation* [Capacity and level of service: Background and documentation]. N.p. Draft.

VEJDIREKTORATET 1999*b*. *Kapacitet og serviceniveau* [Capacity and level of service]. N.p. Draft.

VINCENT, R. A. 1973. Junction priority for public transport. In *Bus Priority—Proceedings of a Symposium held at TRRL, 1972.* (TRRL Report LR 570). Transport and Road Research Laboratory. Crowthorne, Berkshire, p. 91–98.

VINCENT, R. A. & YOUNG, C. P. 1986. Self-optimising traffic signal control using microprocessors—the TRRL 'MOVA' strategy for isolated intersections. *Traffic Engineering and Control*, vol. 27, no. 7/8, p. 385–387.

WATSON, H. 1933. *Street Traffic Flow.* London: Chapman and Hall Ltd.

WEBSTER, F. V. 1958. *Traffic Signal Settings.* (Road Research Technical Paper No. 39). London: Her Majesty's Stationery Office.

WEBSTER, F. V. & COBBE, B. 1966. *Traffic Signals.* (Road Research Technical Paper No. 56). London: Her Majesty's Stationery Office.

WELSH, P. 2001. Saturation flow at junctions—the variable design parameter. *Traffic Engineering and Control*, vol. 42, no. 11, p. 381–385.

WILDERMUTH, B. R. 1962. Average vehicle headways at signalized intersections. *Traffic Engineering*, vol. 32, November, p. 14–16.

WU, N. 1992. Wartezeiten an festzeitgesteuerten lichtsignalanlagen unter zeitlich veränderlichen (instationären) verkehrsbedingungen [Delays at pretimed traffic signals under time-dependent (nonstationary) traffic conditions]. *Straßenverkehrstechnik* 3/92, p. 147–153.

YAGAR, S. 1977. Minimizing delays for transient demands with application to signalized road junctions. *Transportation Research*, vol. 11, no. 1, p. 53–62.

APPENDICES

A INTERSECTION BASIC-1

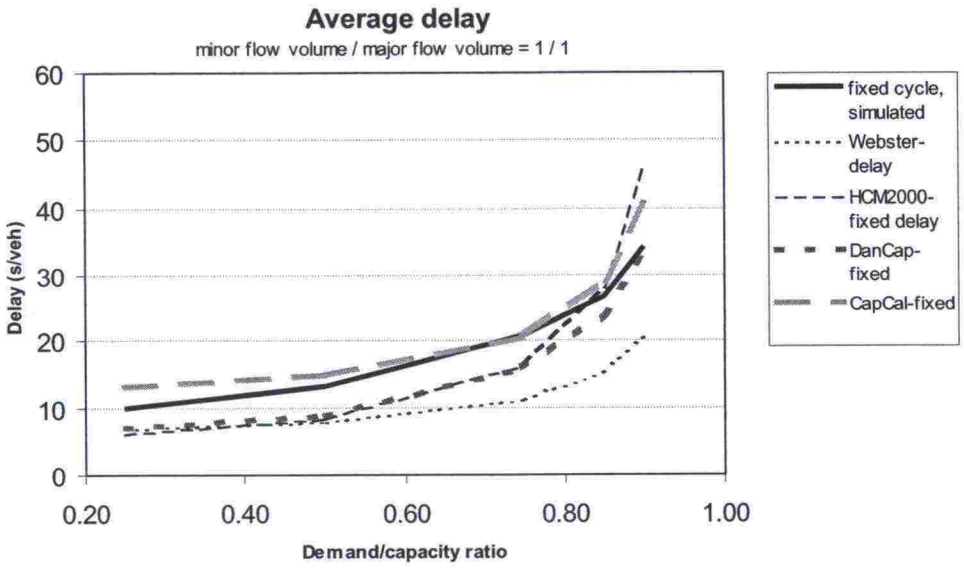


Figure A.1: Delays in pretimed intersection Basic-1 with minor/major flow ratio 1

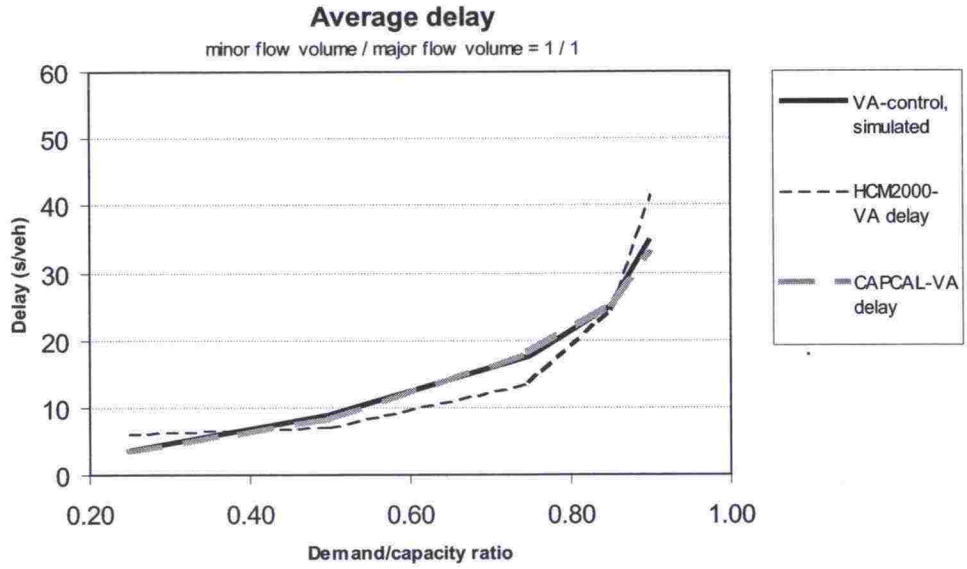


Figure A.2: Delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1

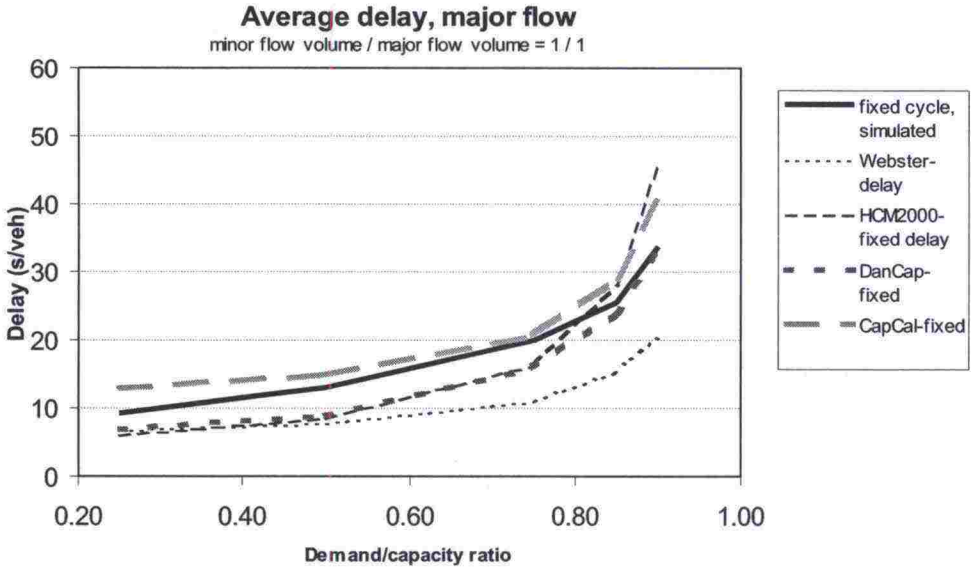


Figure A.3: Major flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1

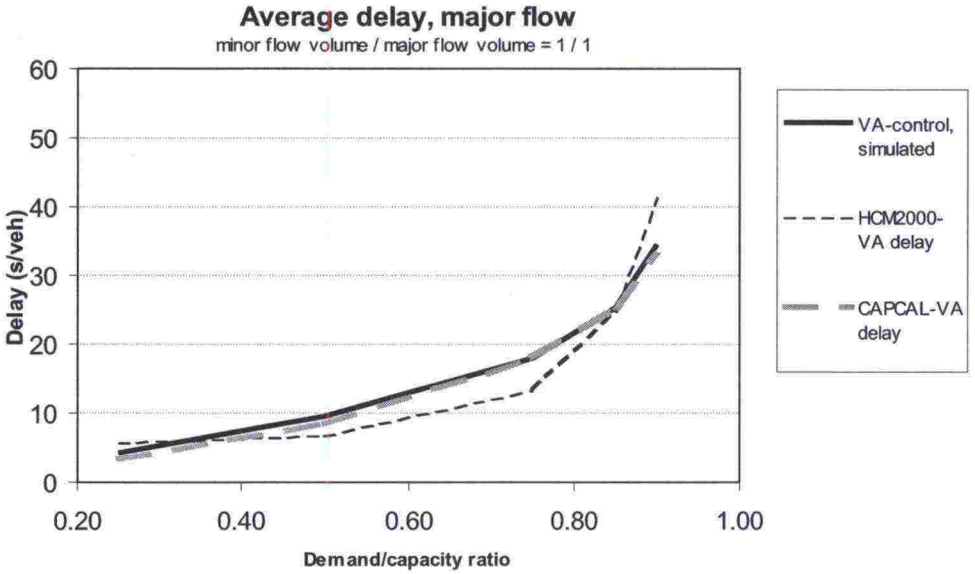


Figure A.4: Major flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1

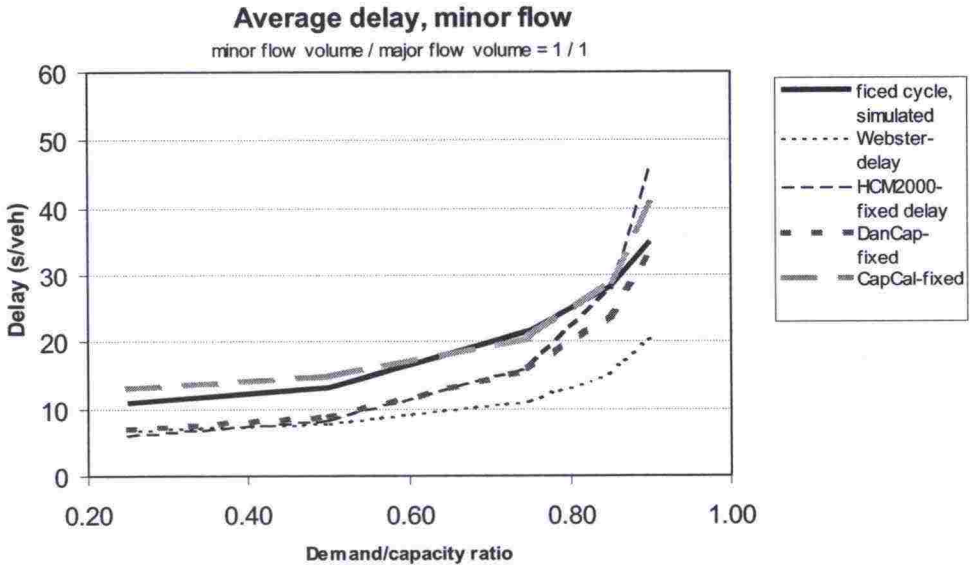


Figure A.5: Minor flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1

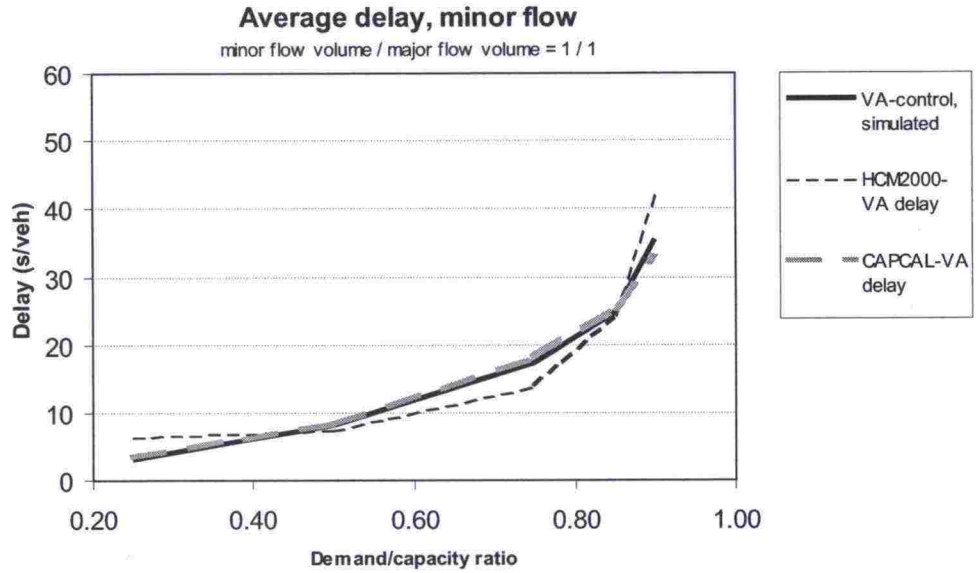


Figure A.6: Minor flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1

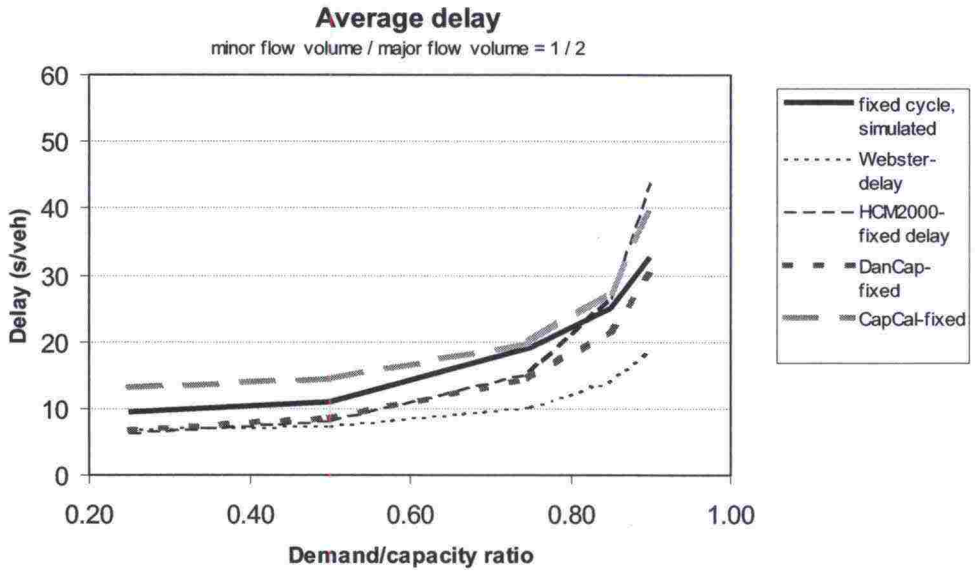


Figure A.7: Delays in pretimed intersection Basic-1 with minor/major flow ratio 1/2

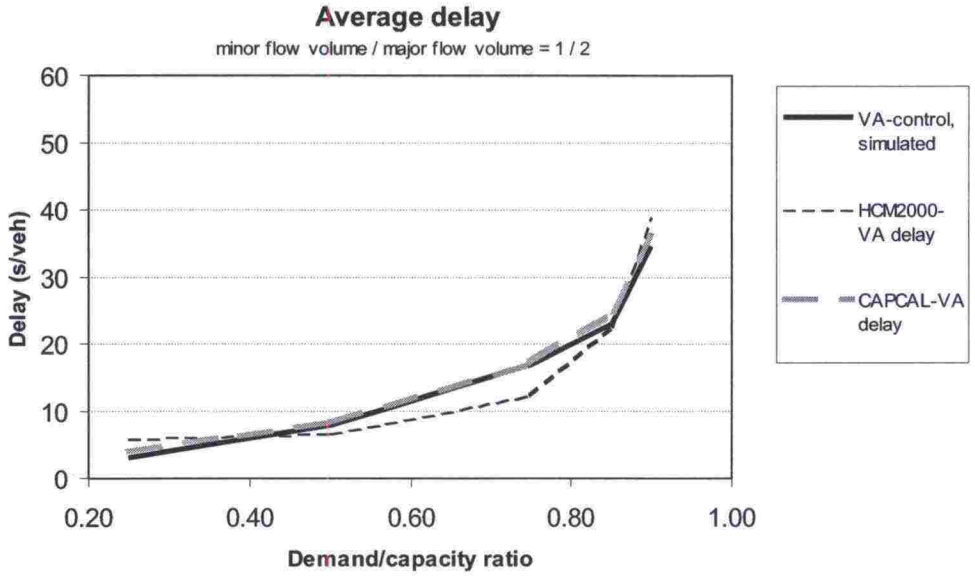


Figure A.8: Delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/2

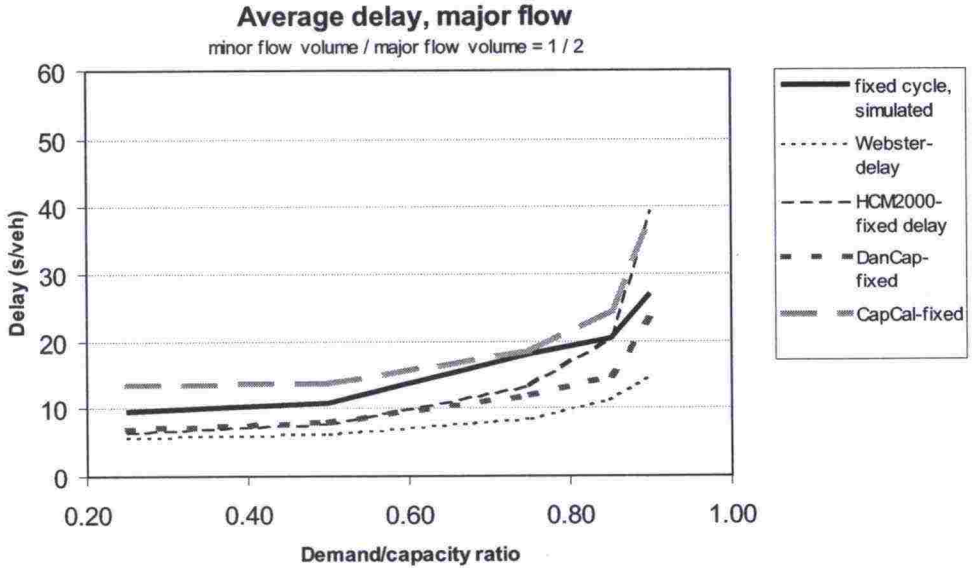


Figure A.9: Major flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1/2

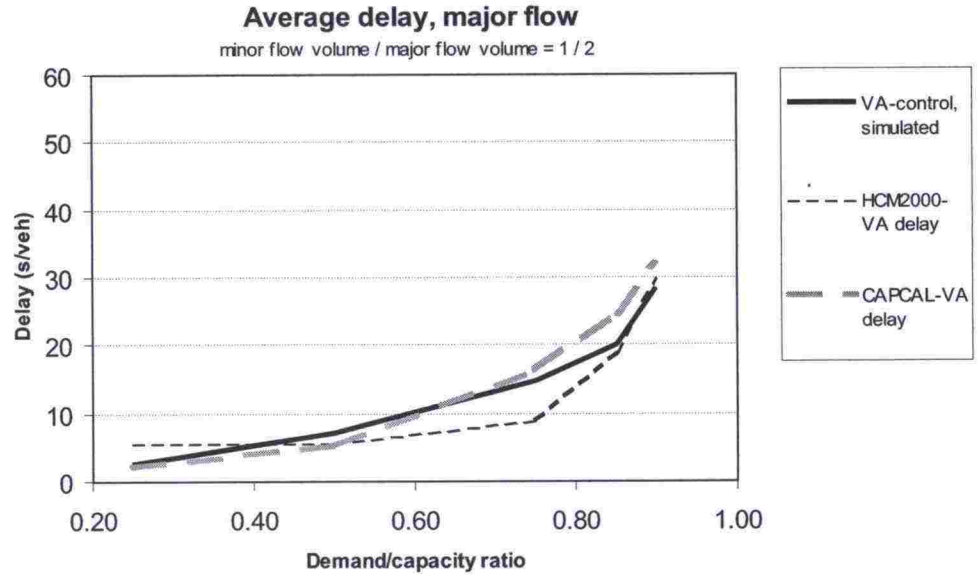


Figure A.10: Major flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/2

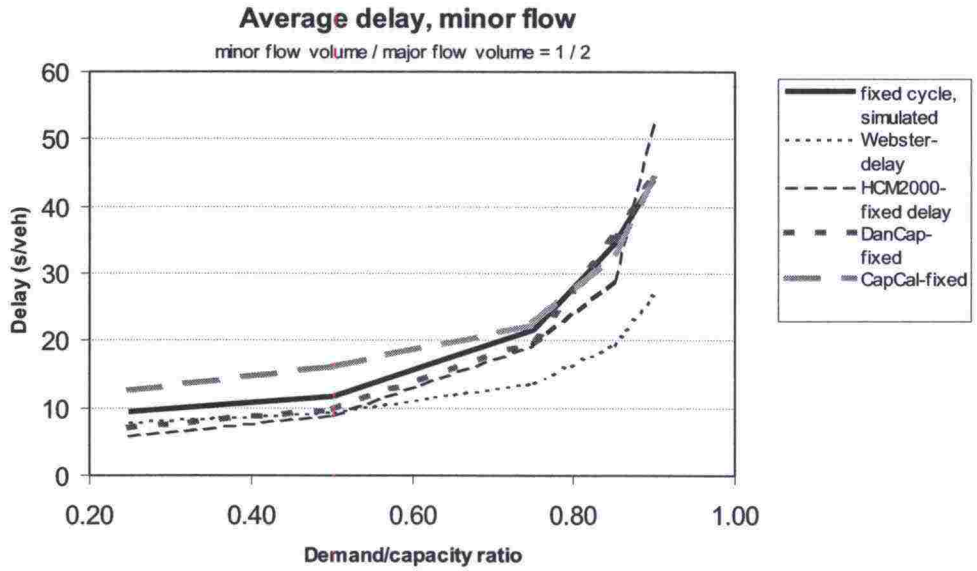


Figure A.11: Minor flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1/2

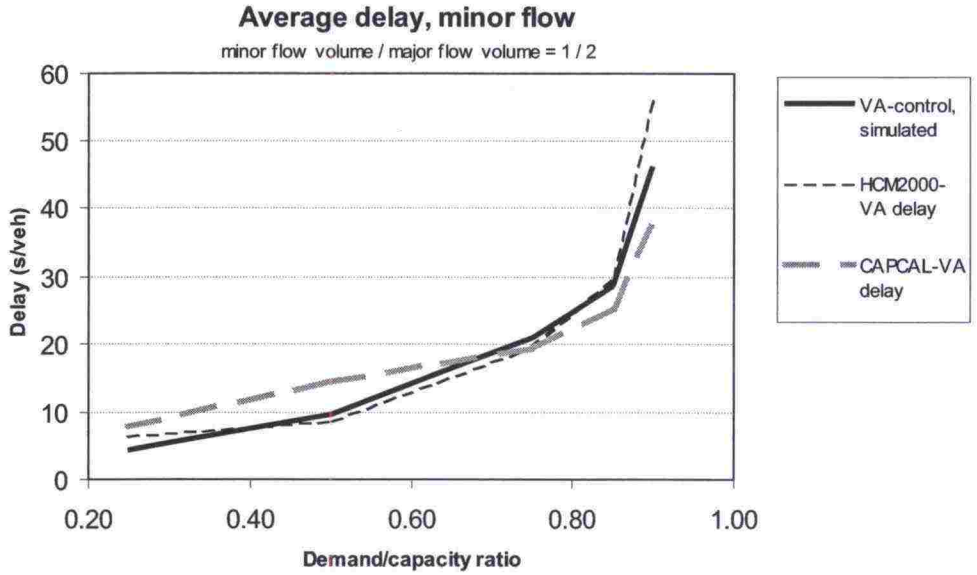


Figure A.12: Minor flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/2

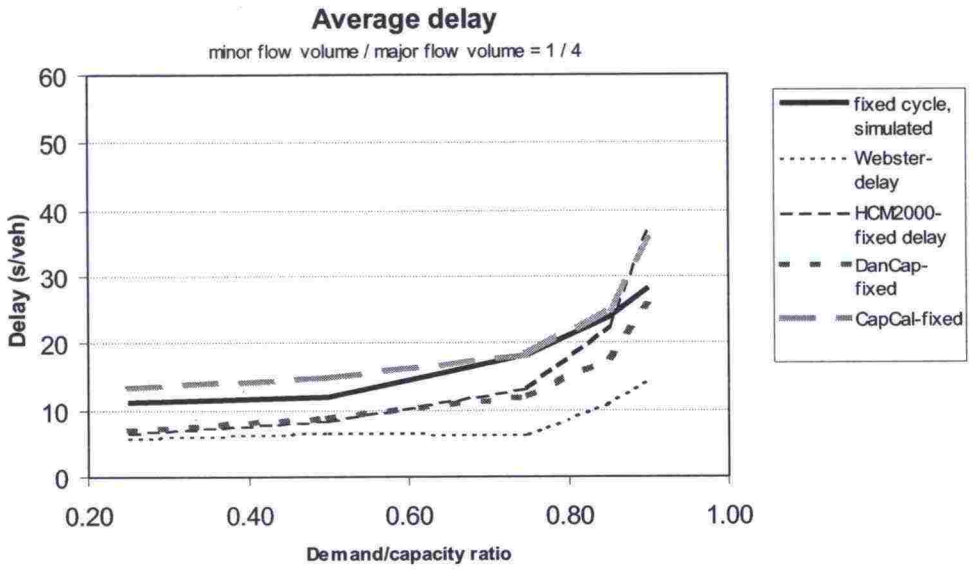


Figure A.13: Delays in pretimed intersection Basic-1 with minor/major flow ratio 1/4

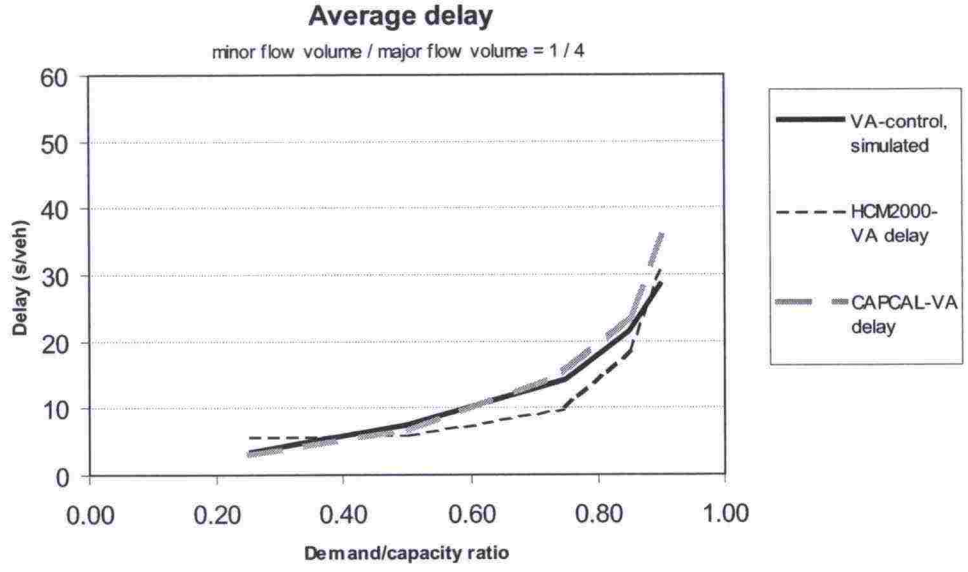


Figure A.14: Delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/4

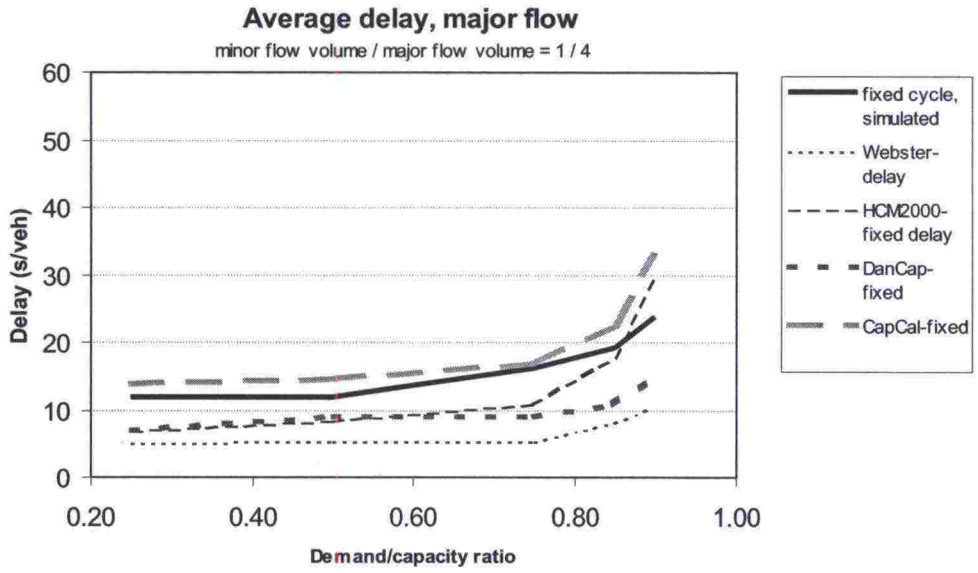


Figure A.15: Major flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1/4

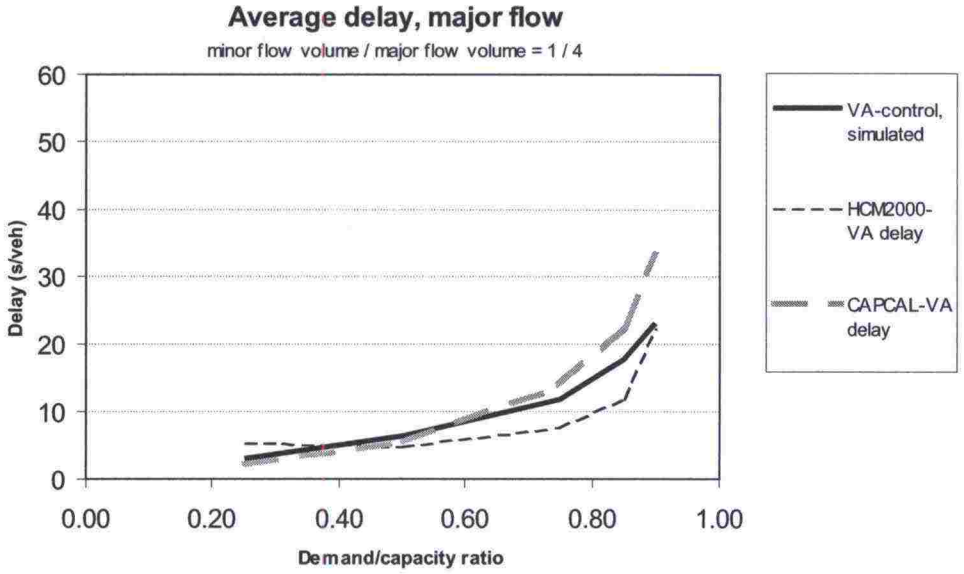


Figure A.16: Major flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/4

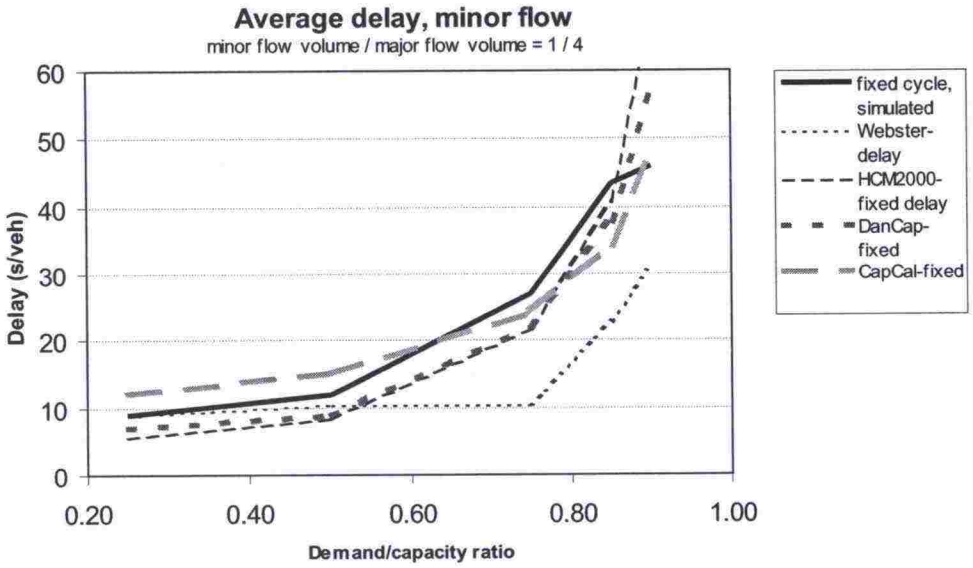


Figure A.17: Minor flow delays in pretimed intersection Basic-1 with minor/major flow ratio 1/4

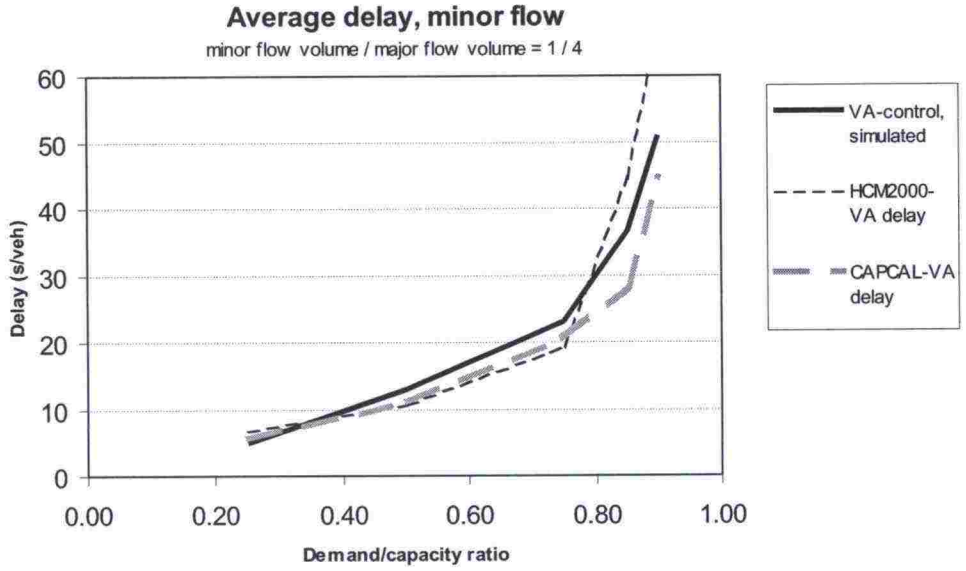


Figure A.18: Minor flow delays in traffic-responsive intersection Basic-1 with minor/major flow ratio 1/4

B INTERSECTION BASIC-2 WITH 10 % LEFT TURNERS

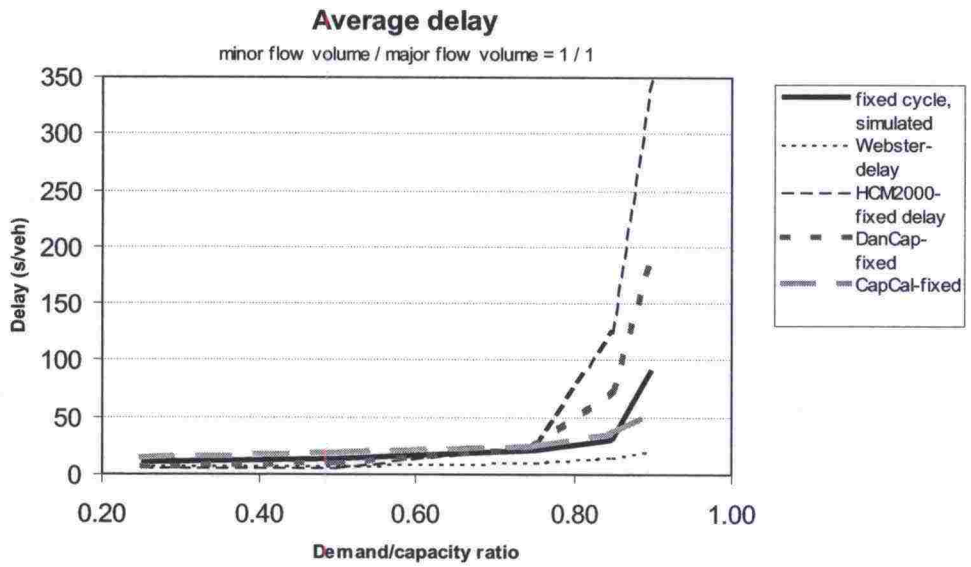


Figure B.1: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

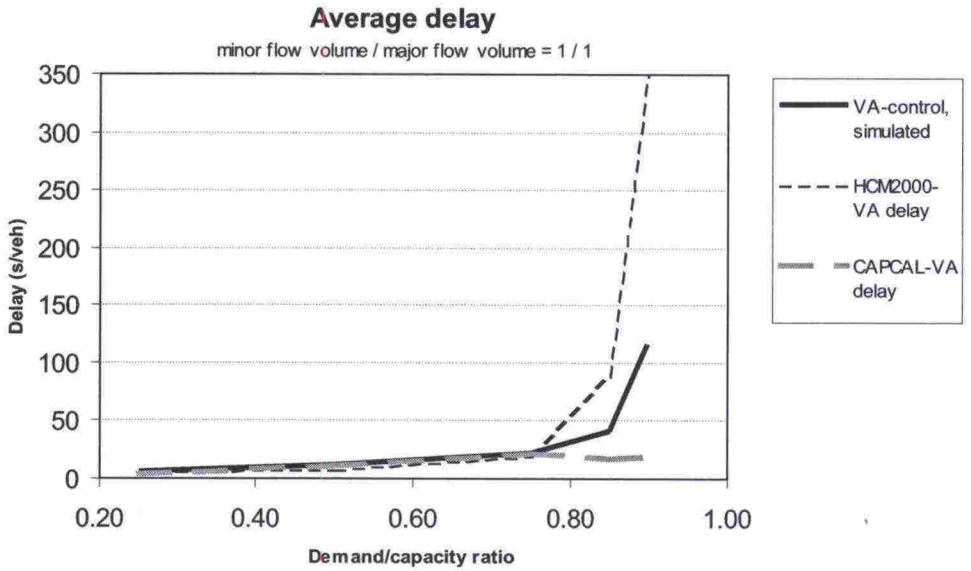


Figure B.2: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

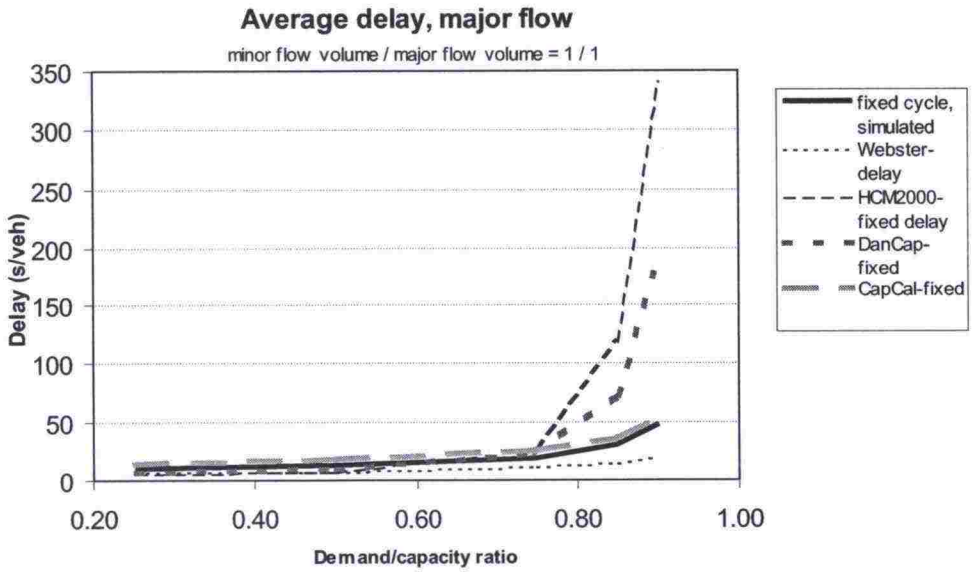


Figure B.3: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

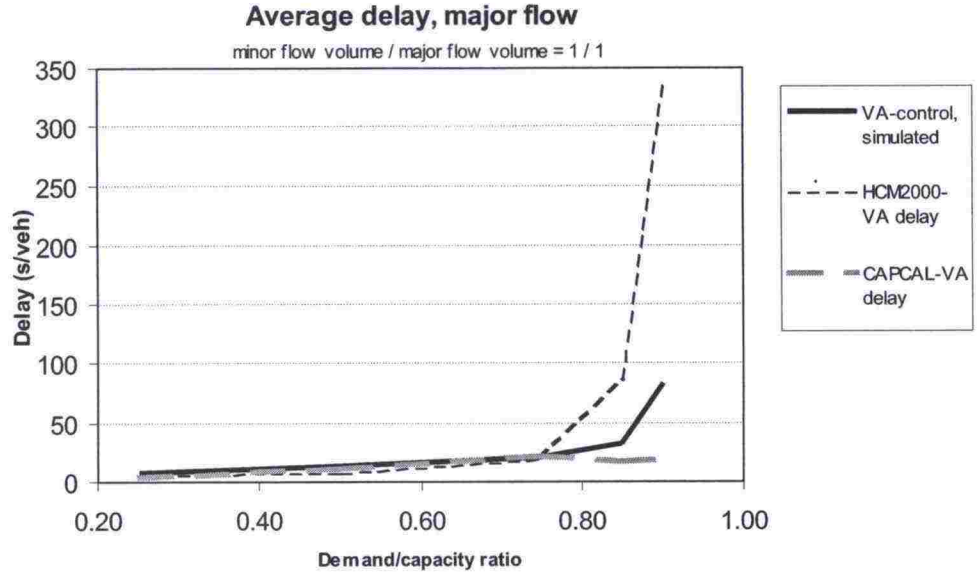


Figure B.4: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

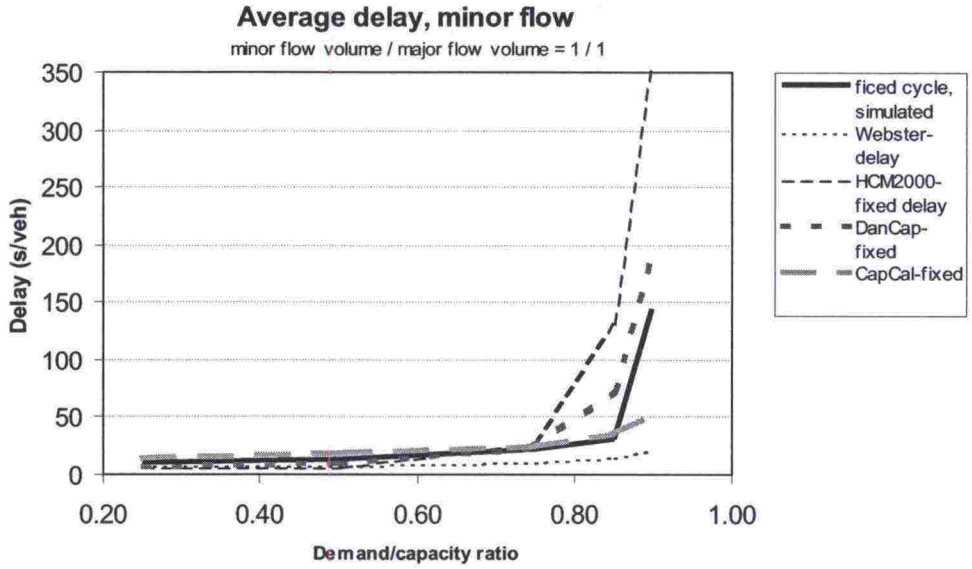


Figure B.5: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

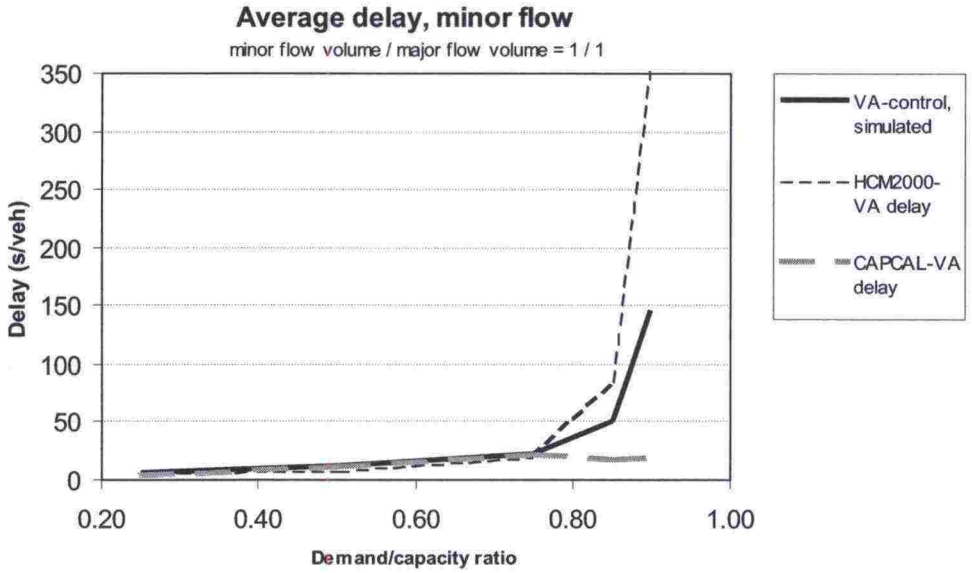


Figure B.6: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 10 % left turners

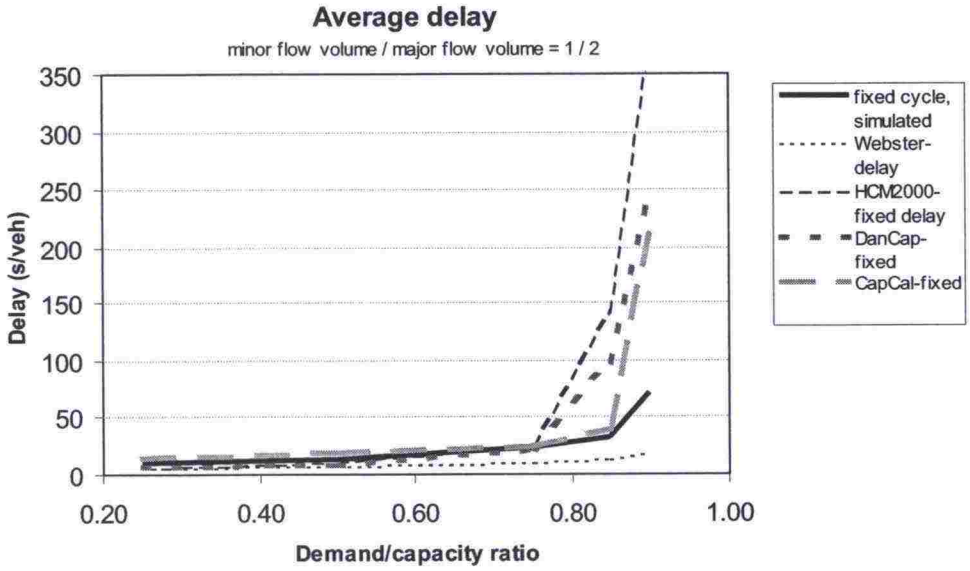


Figure B.7: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

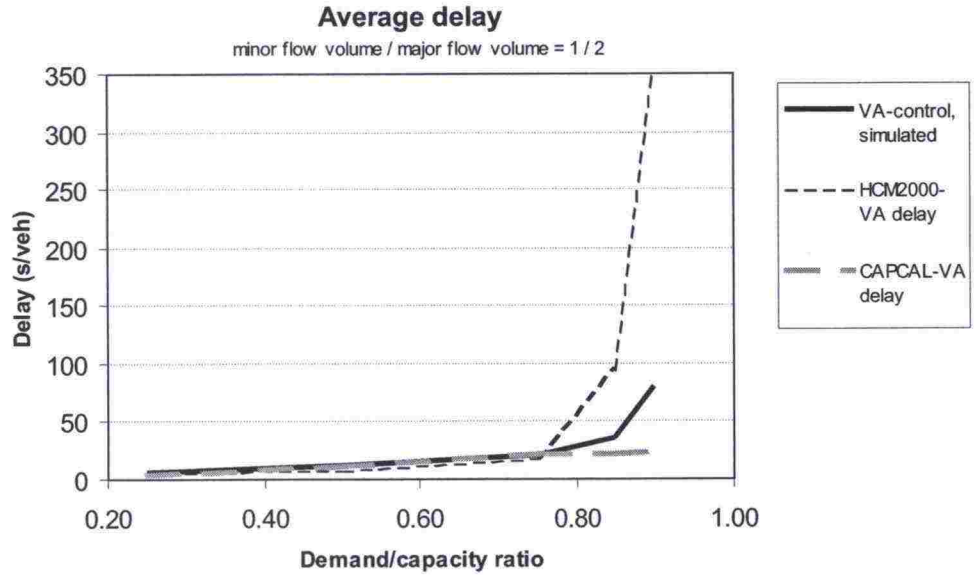


Figure B.8: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

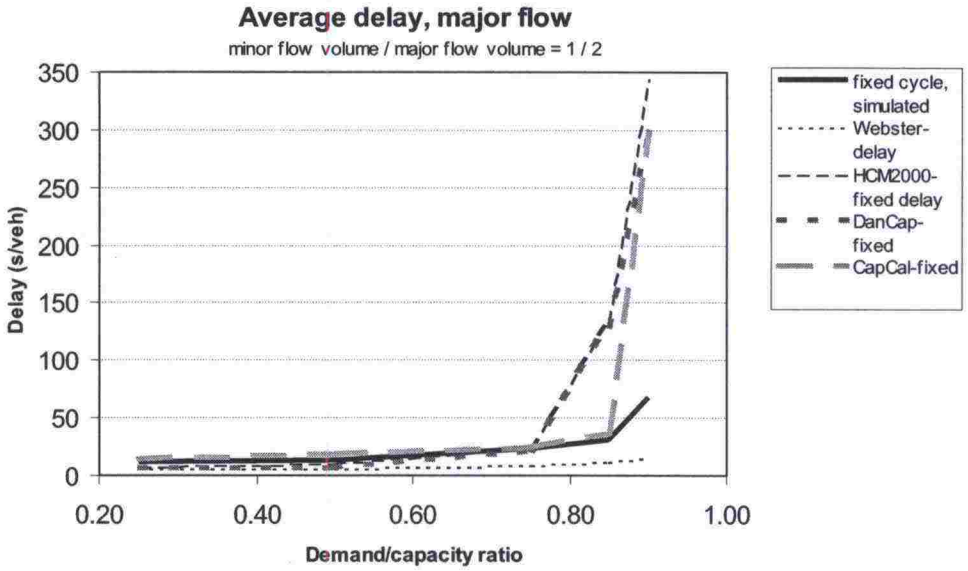


Figure B.9: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

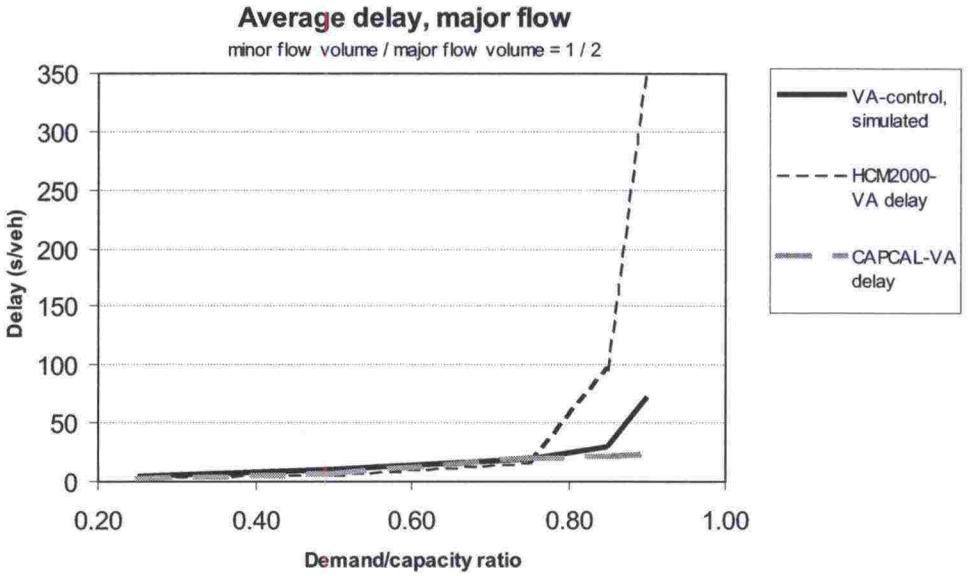


Figure B.10: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

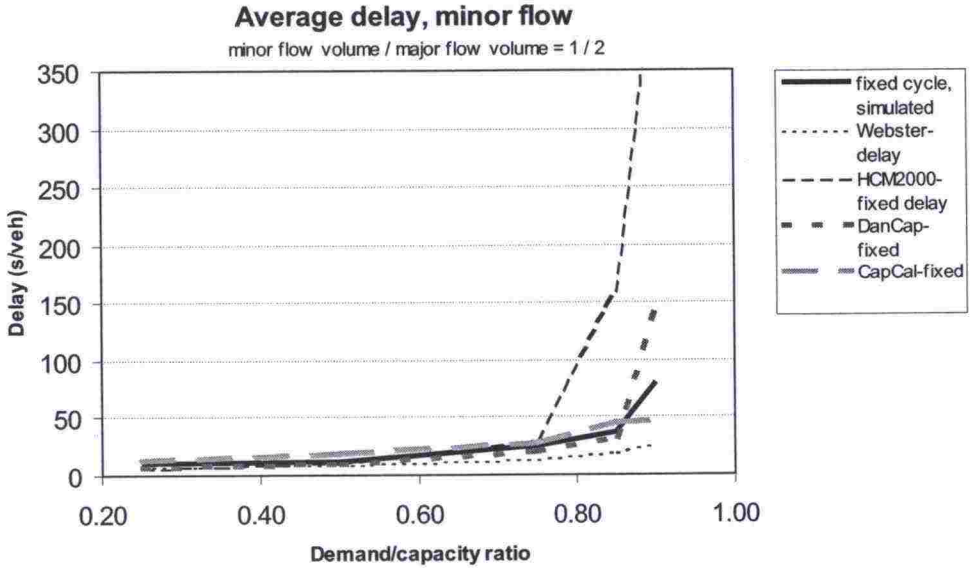


Figure B.11: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

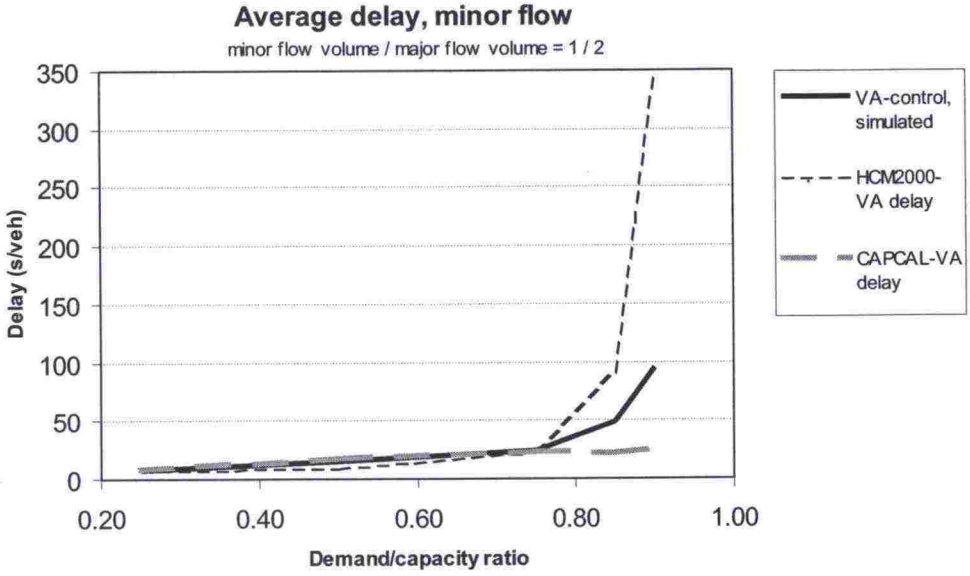


Figure B.12: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 10 % left turners

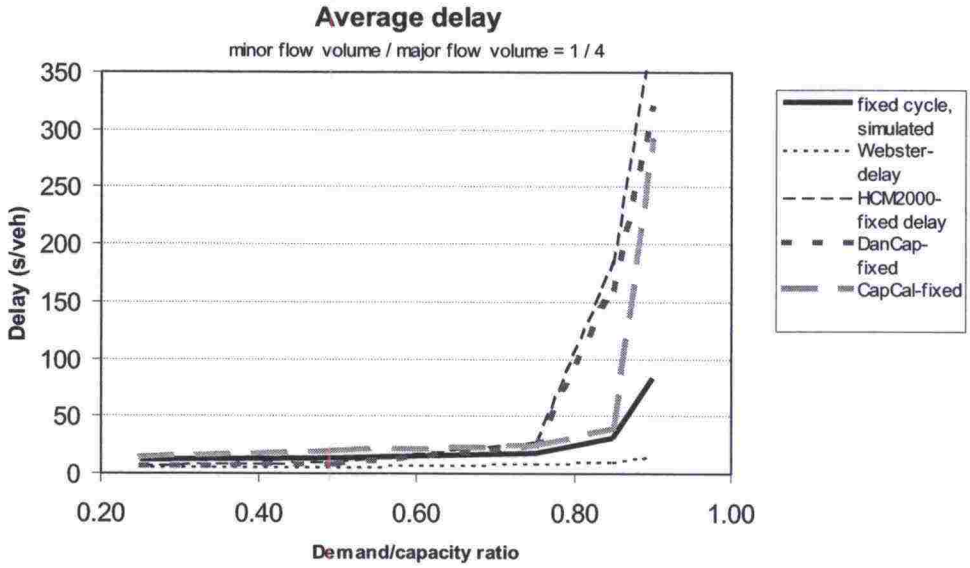


Figure B.13: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

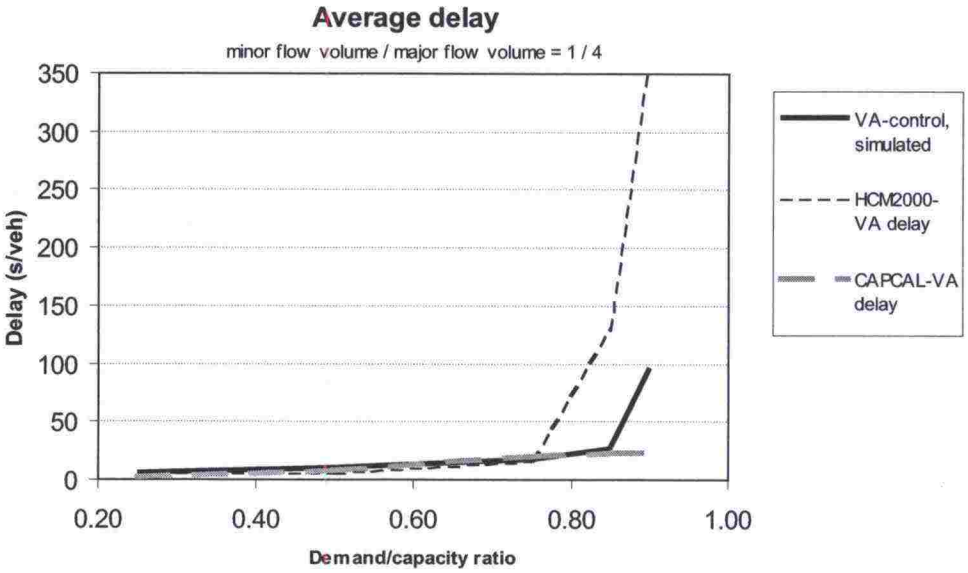


Figure B.14: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

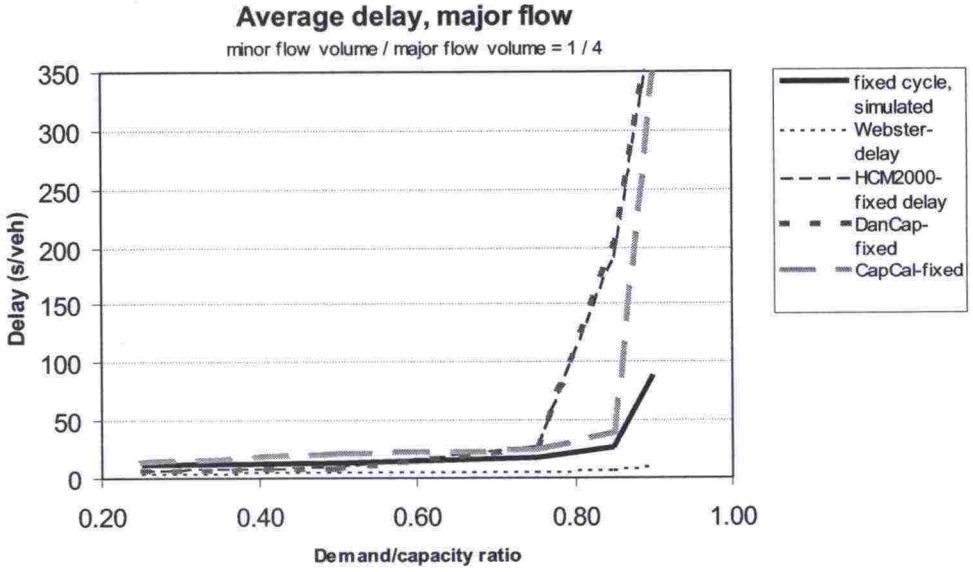


Figure B.15: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

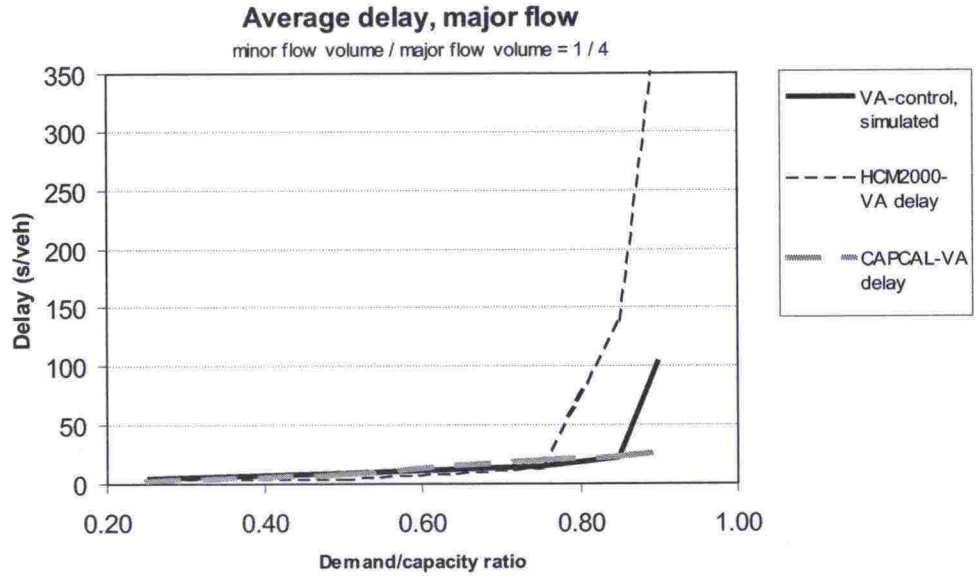


Figure B.16: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

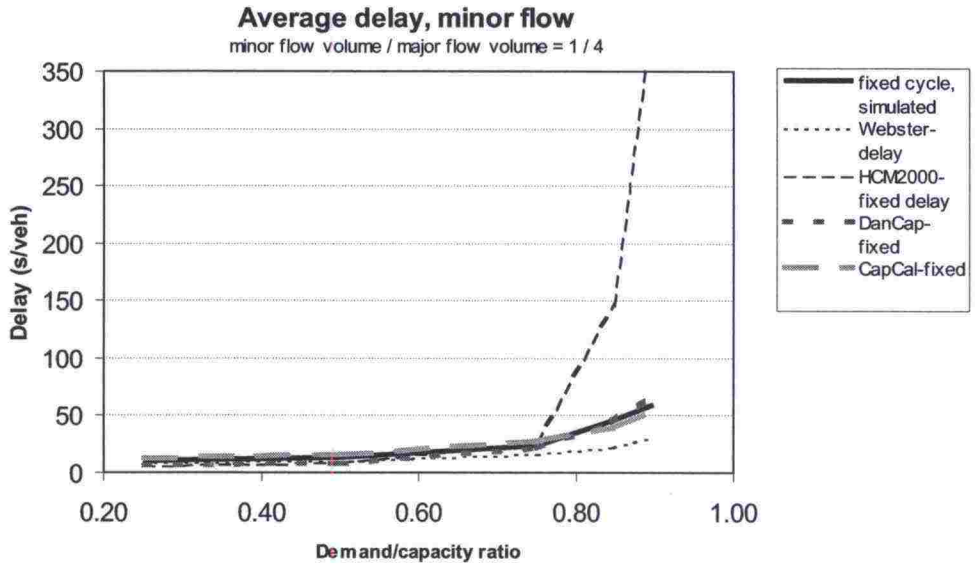


Figure B.17: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

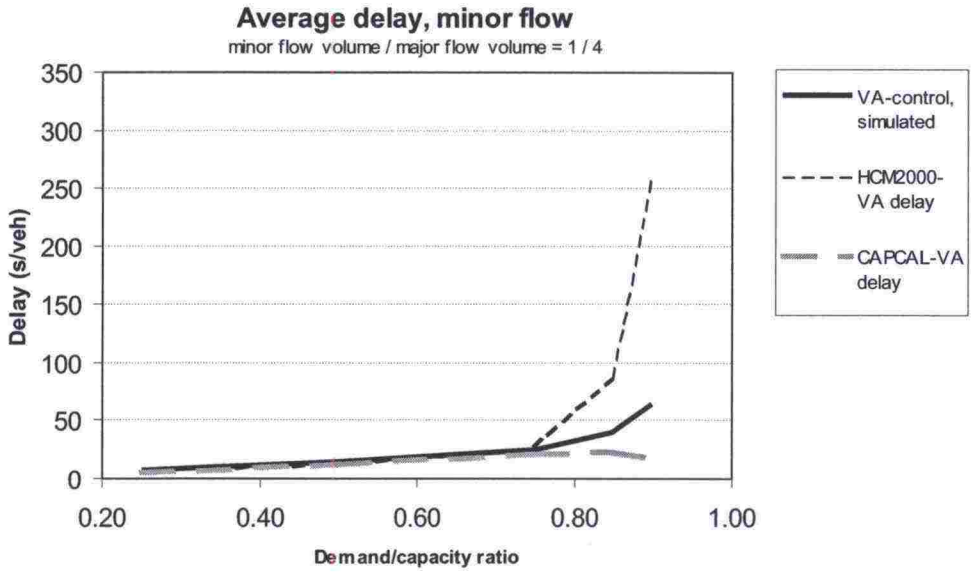


Figure B.18: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 10 % left turners

C INTERSECTION BASIC-2 WITH 25 % LEFT TURNERS

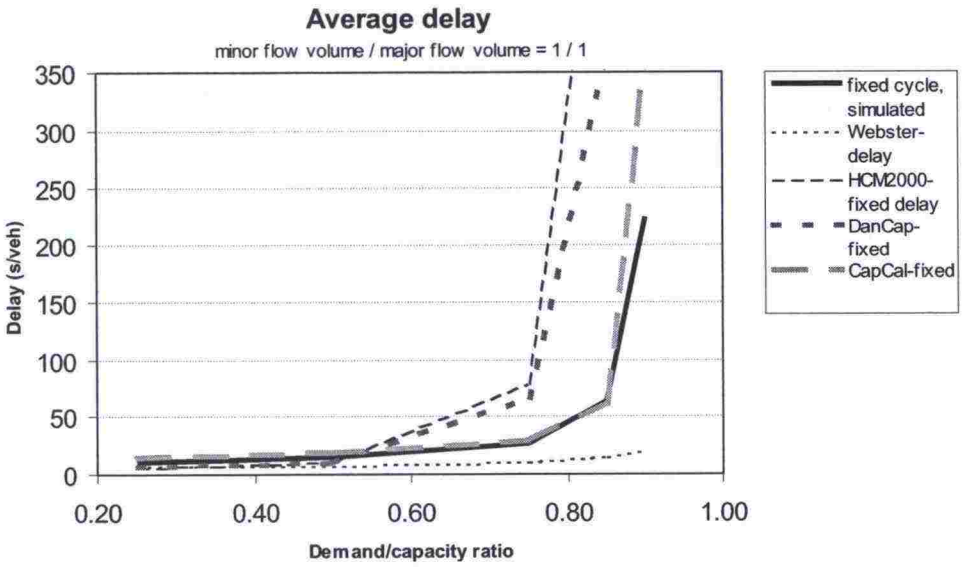


Figure C.1: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

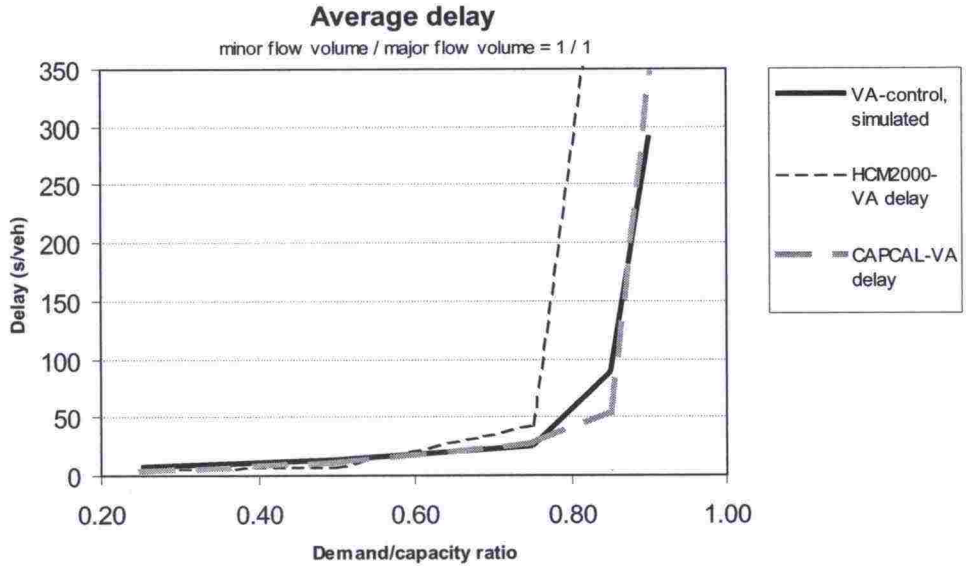


Figure C.2: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

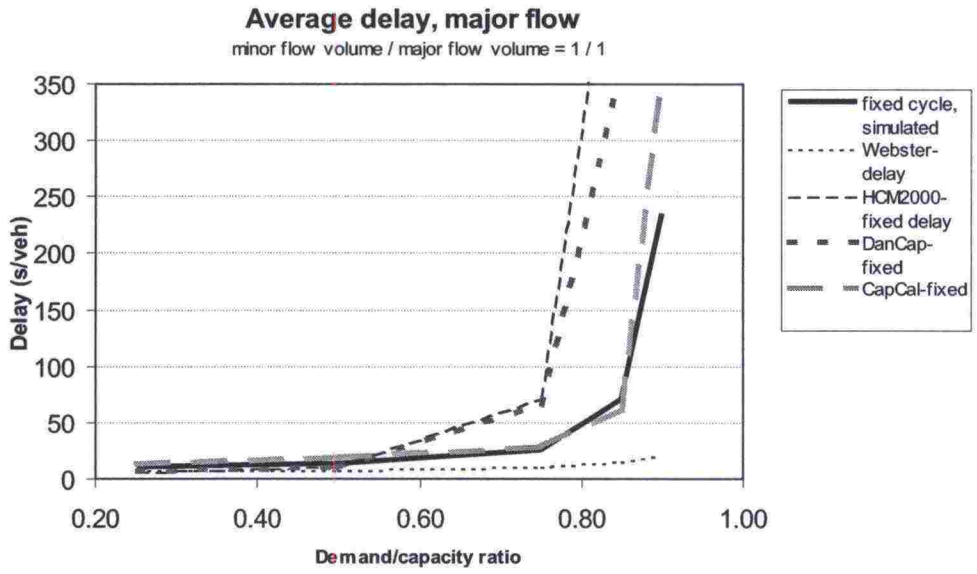


Figure C.3: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

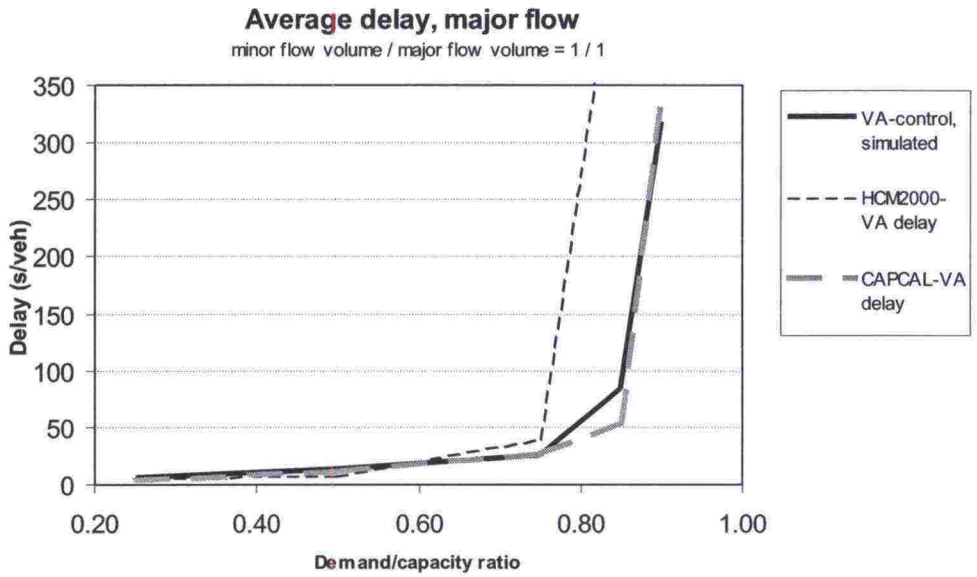


Figure C.4: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

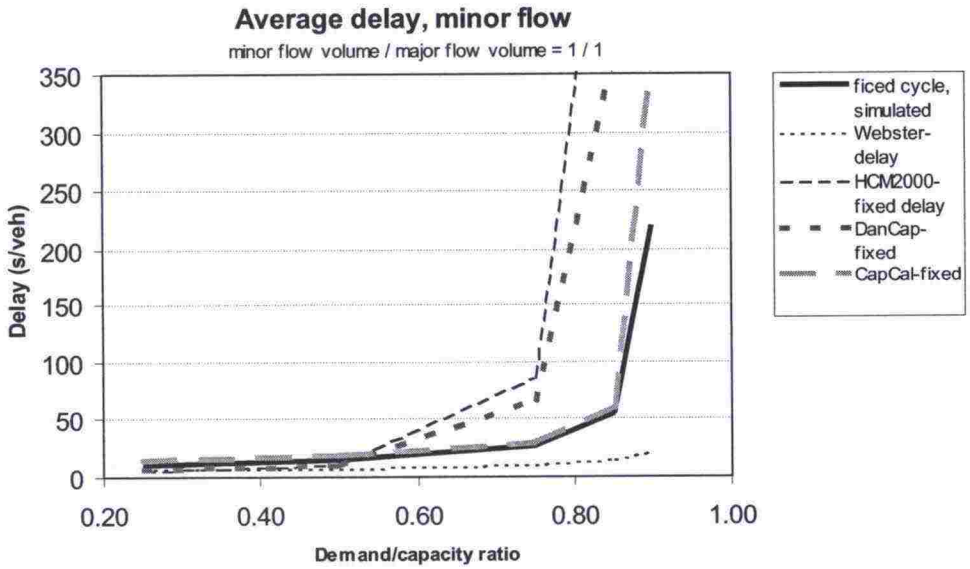


Figure C.5: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

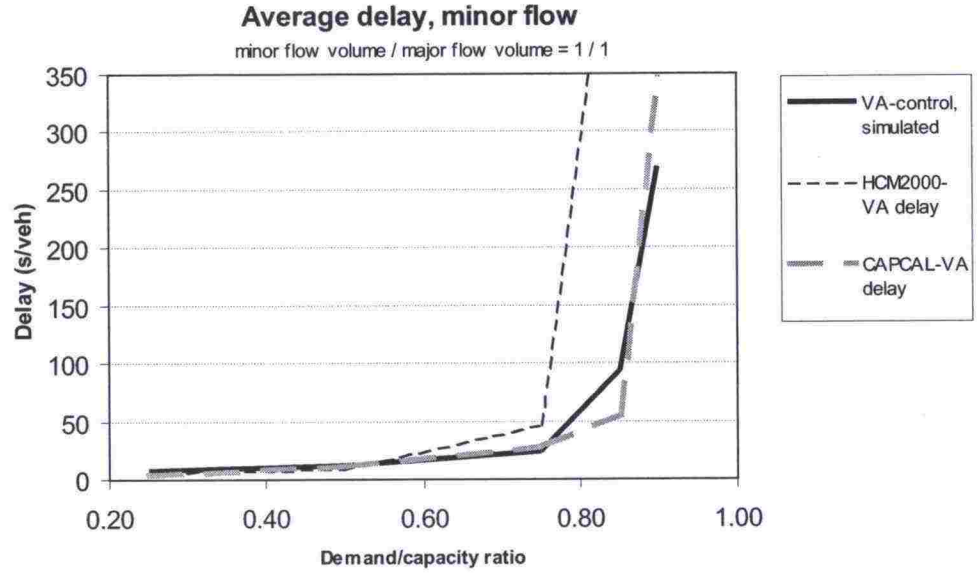


Figure C.6: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1 and 25 % left turners

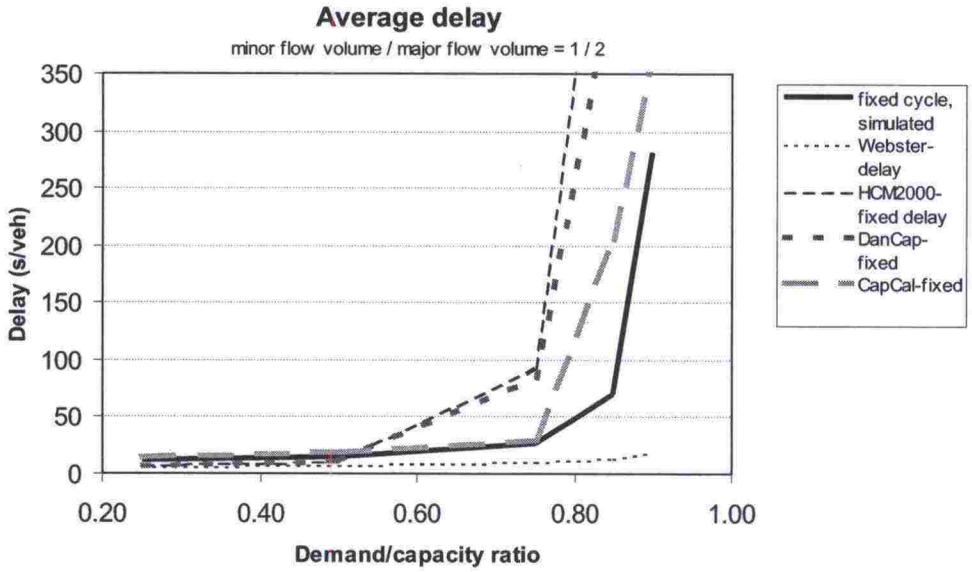


Figure C.7: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

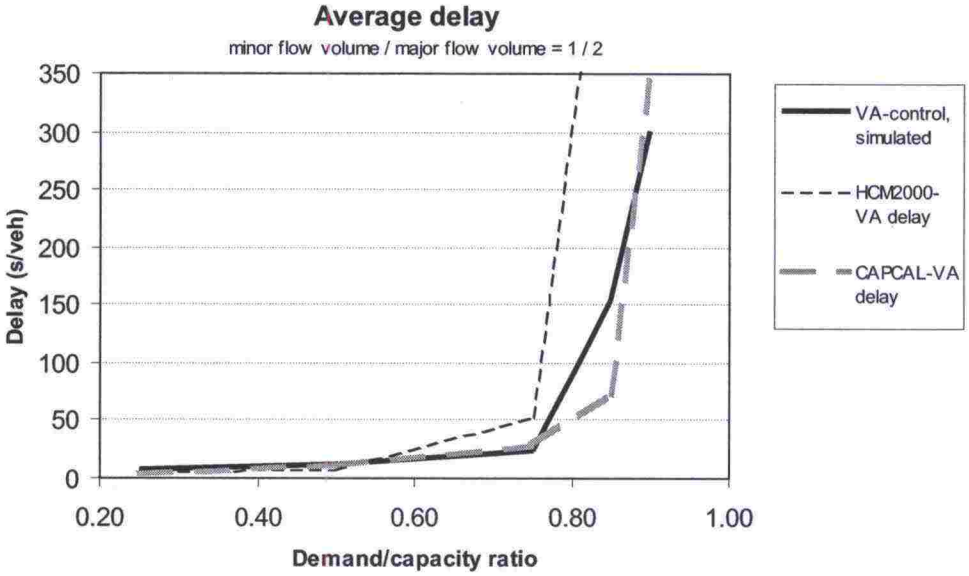


Figure C.8: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

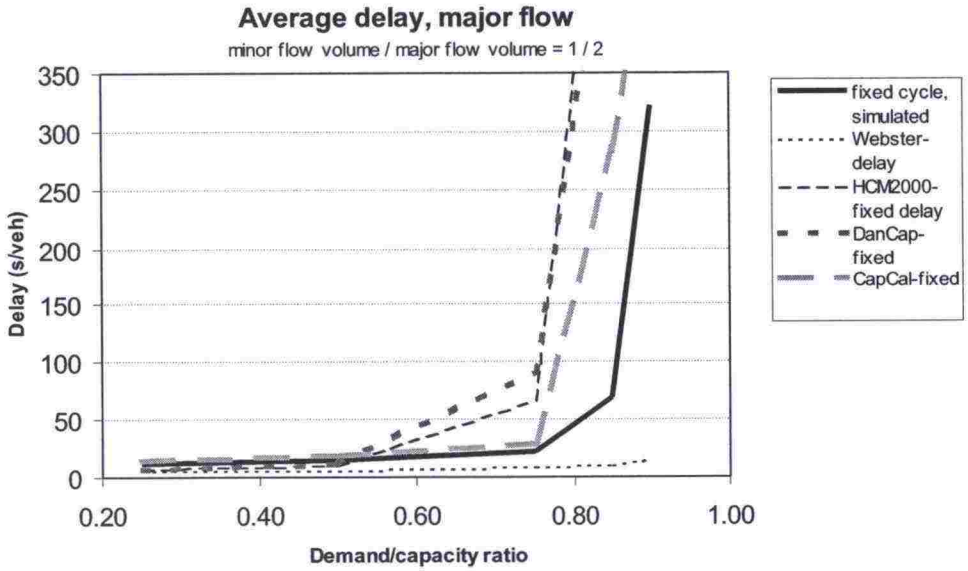


Figure C.9: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

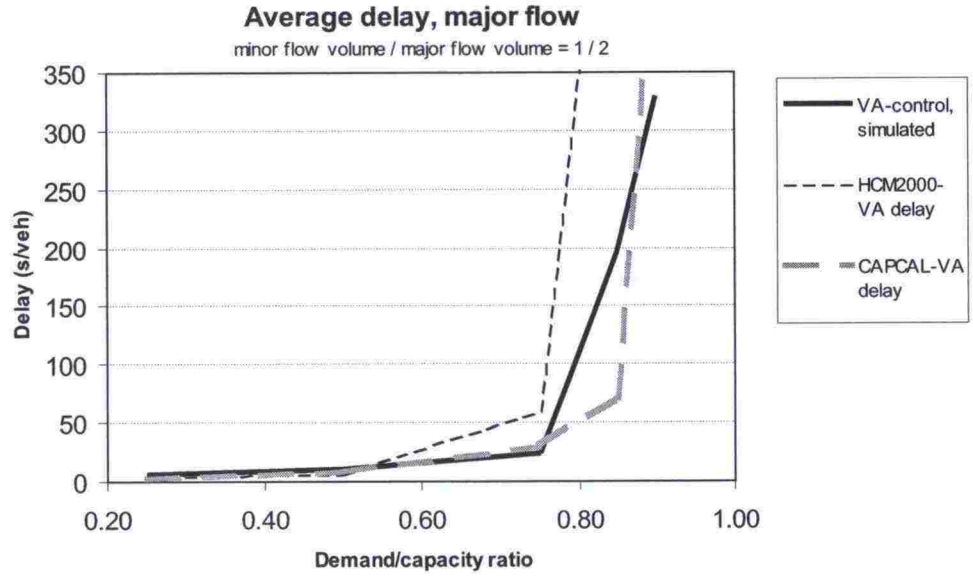


Figure C.10: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

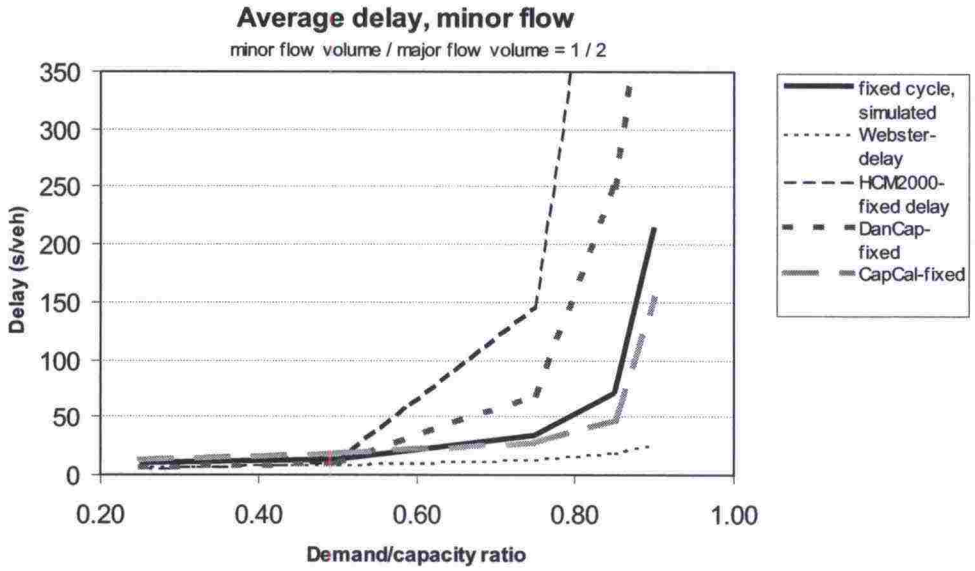


Figure C.11: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

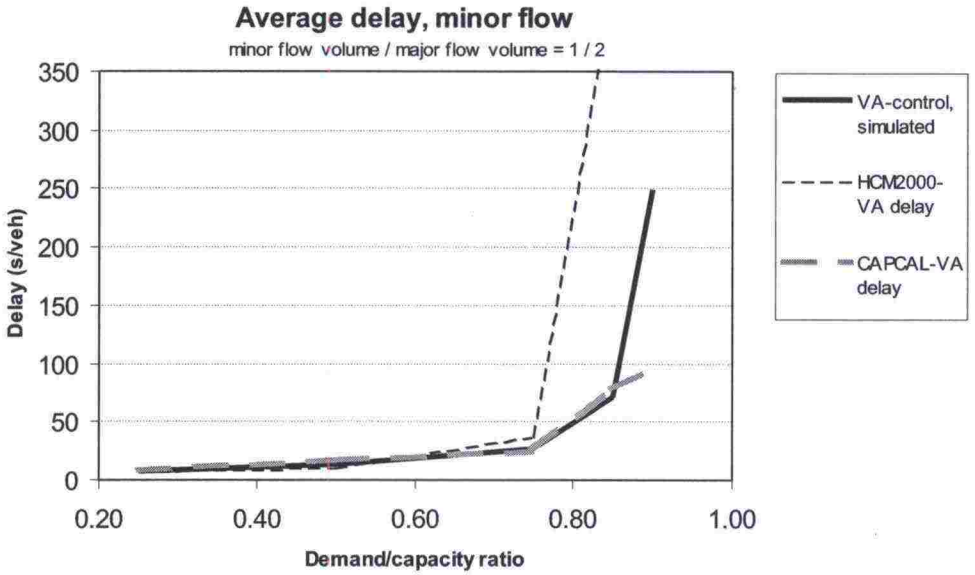


Figure C.12: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/2 and 25 % left turners

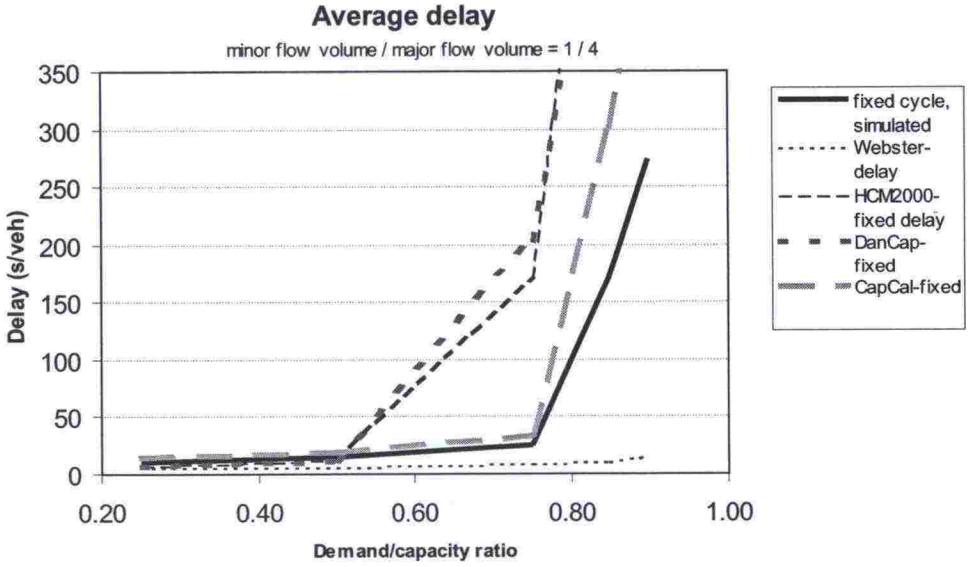


Figure C.13: Delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

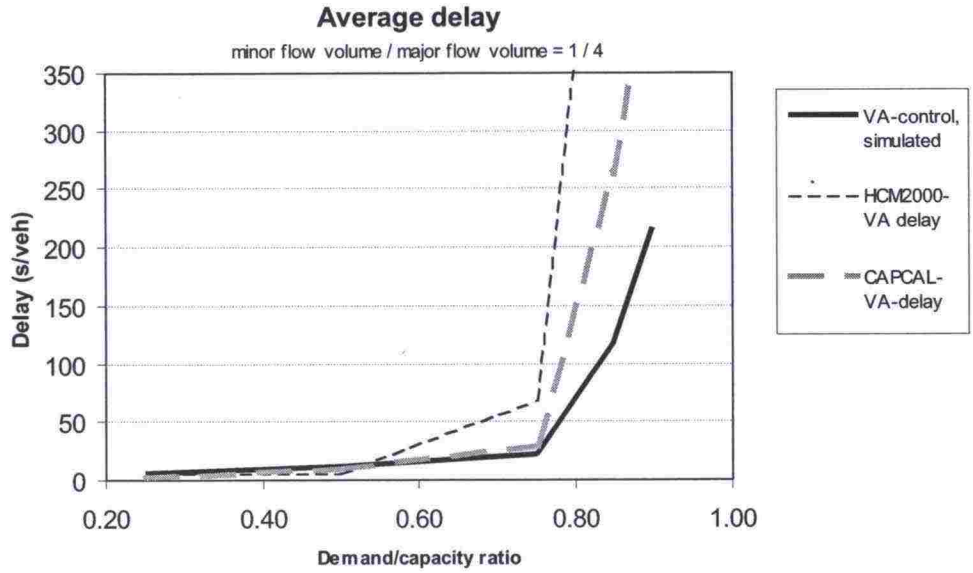


Figure C.14: Delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

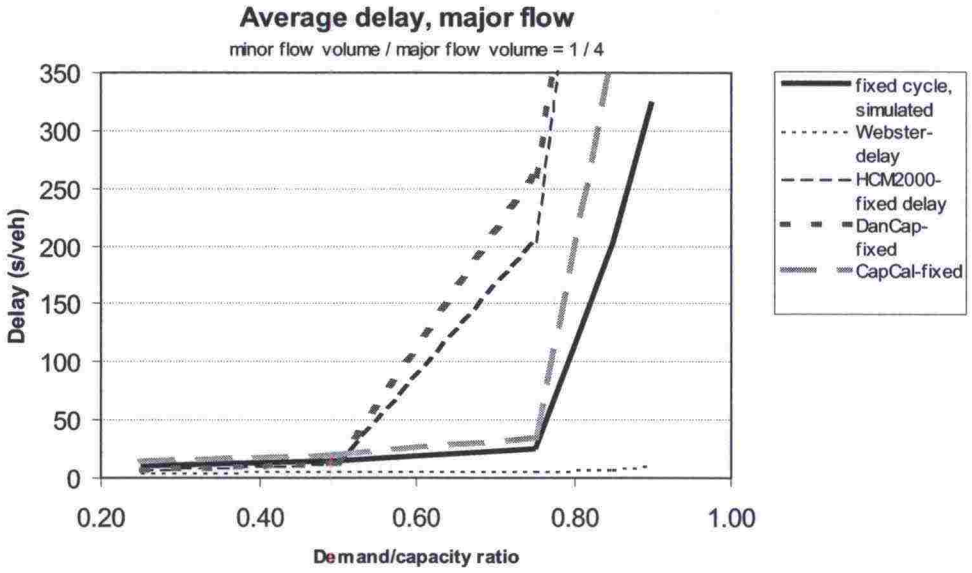


Figure C.15: Major flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

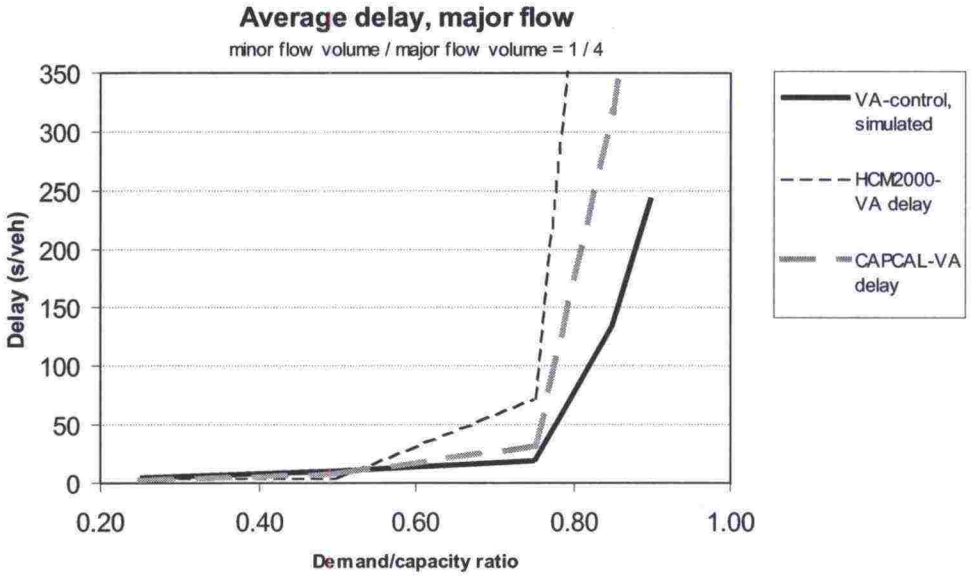


Figure C.16: Major flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

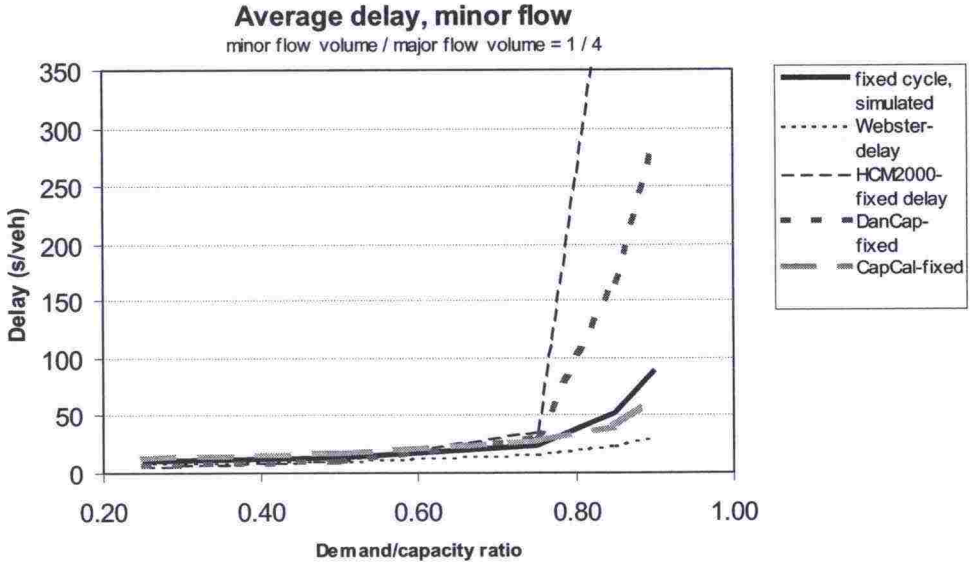


Figure C.17: Minor flow delays in pretimed intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

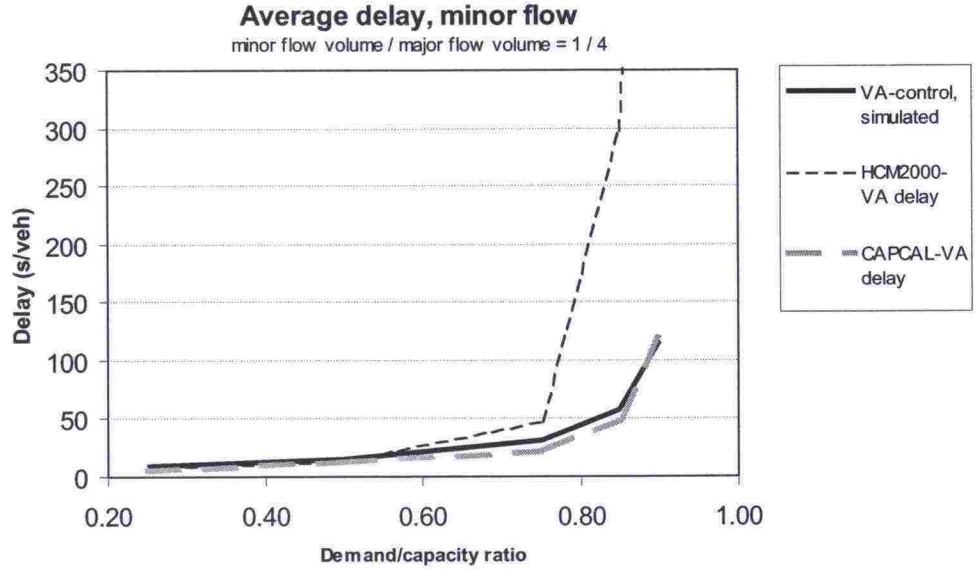


Figure C.18: Minor flow delays in traffic-responsive intersection Basic-2 with minor/major flow ratio 1/4 and 25 % left turners

D INTERSECTION HCM-1

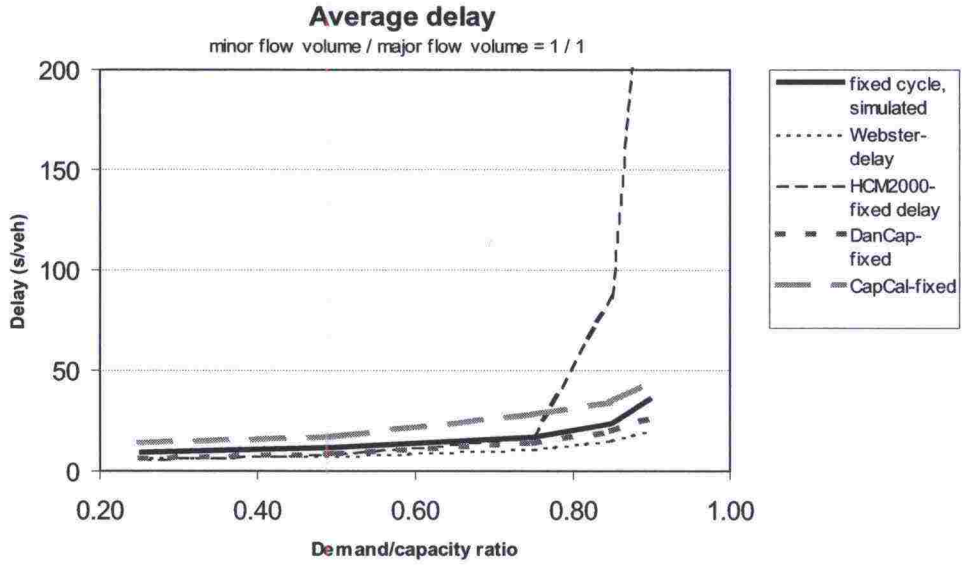


Figure D.1: Delays in pretimed intersection HCM-1 with minor/major flow ratio 1

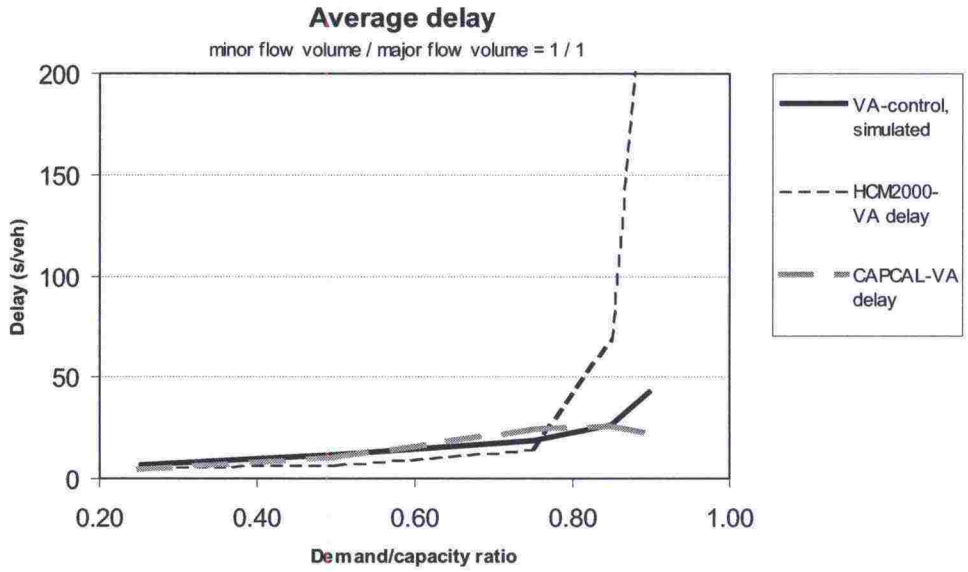


Figure D.2: Delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1

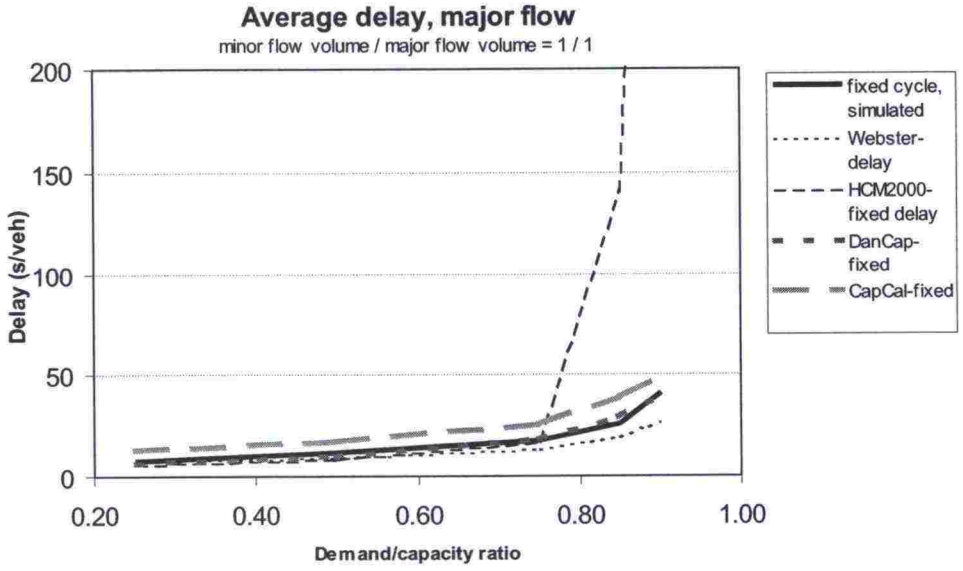


Figure D.3: Major flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1

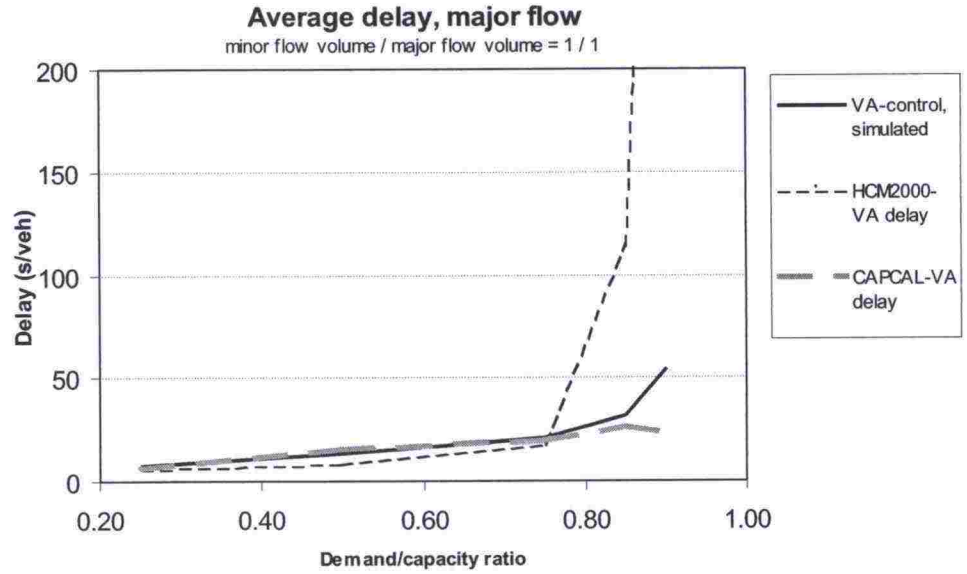


Figure D.4: Major flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1

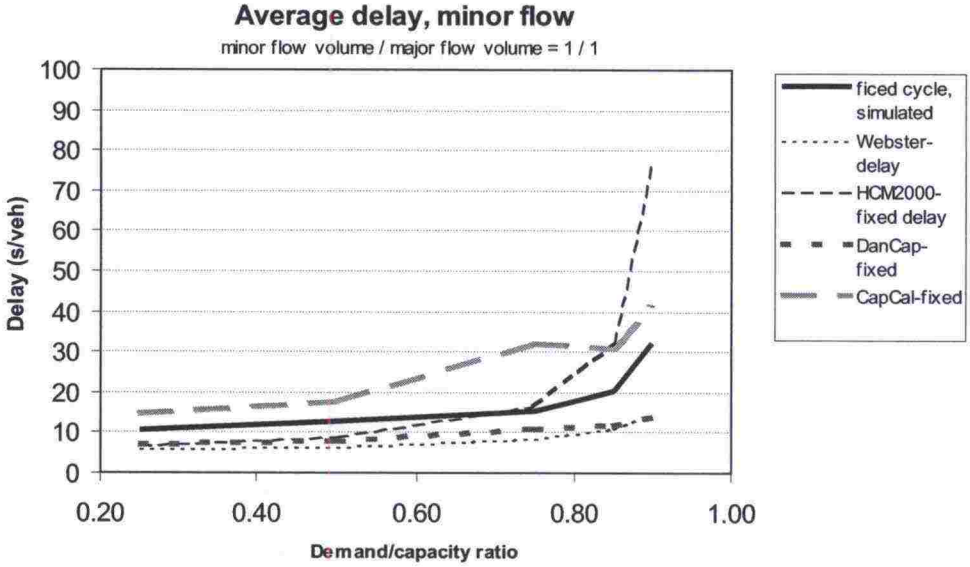


Figure D.5: Minor flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1

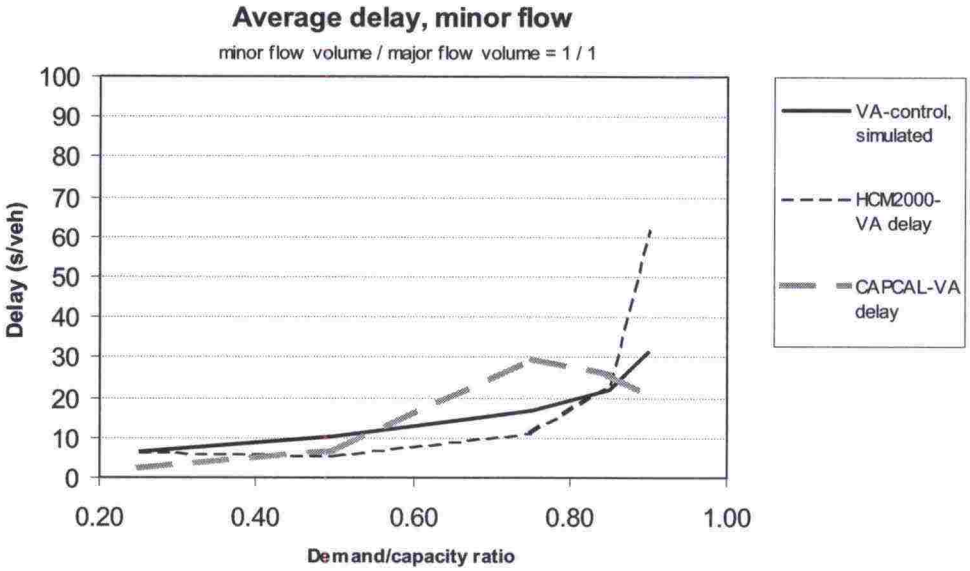


Figure D.6: Minor flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1

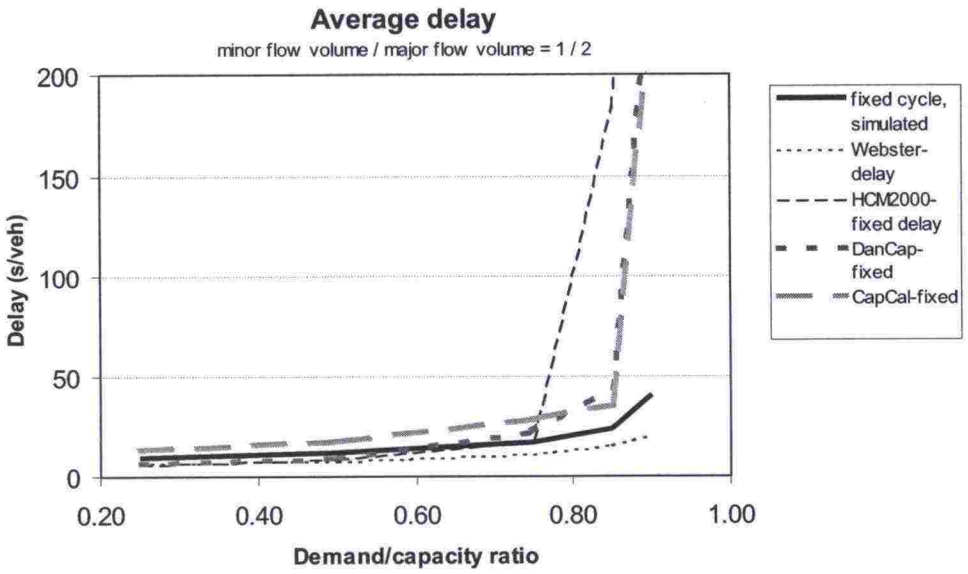


Figure D.7: Delays in pretimed intersection HCM-1 with minor/major flow ratio 1/2

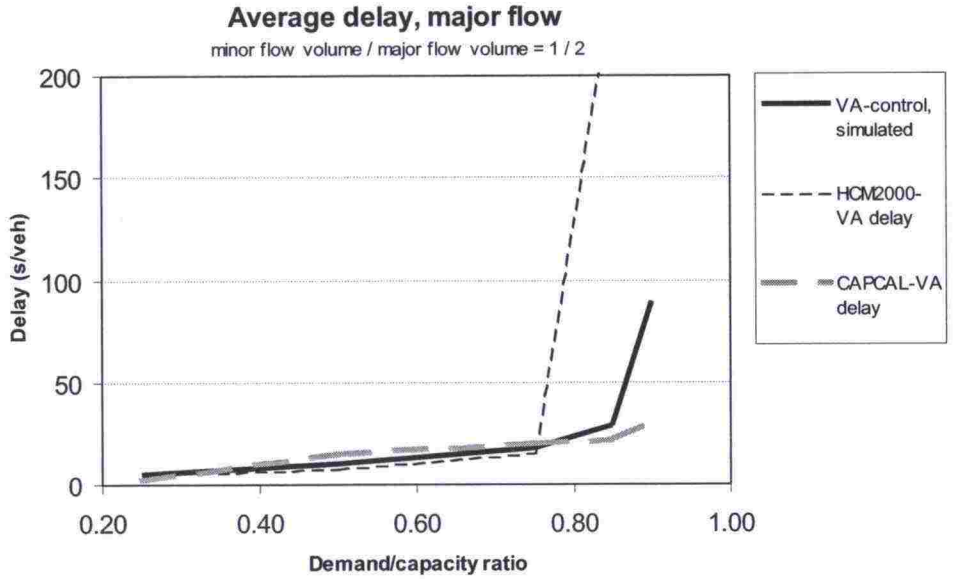


Figure D.8: Delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/2

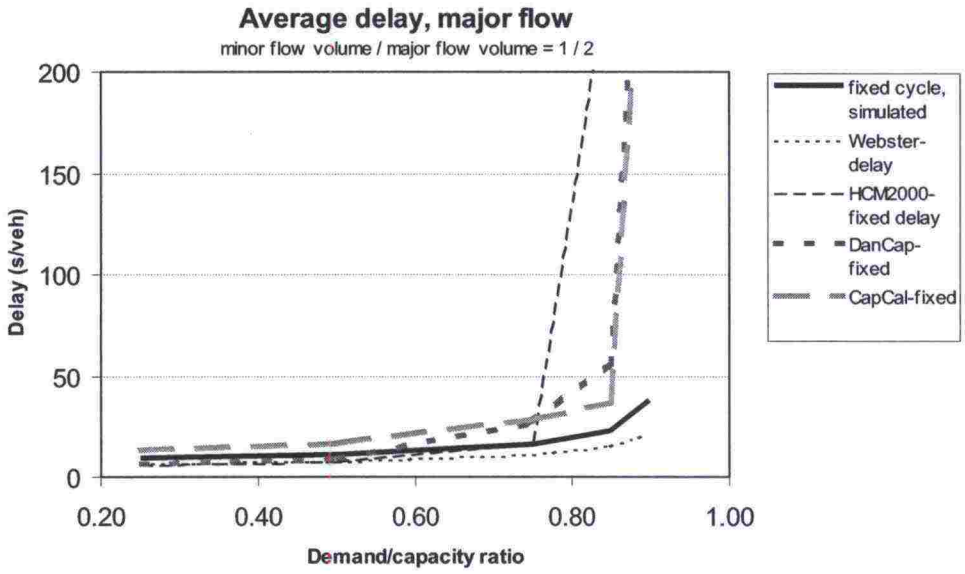


Figure D.9: Major flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1/2

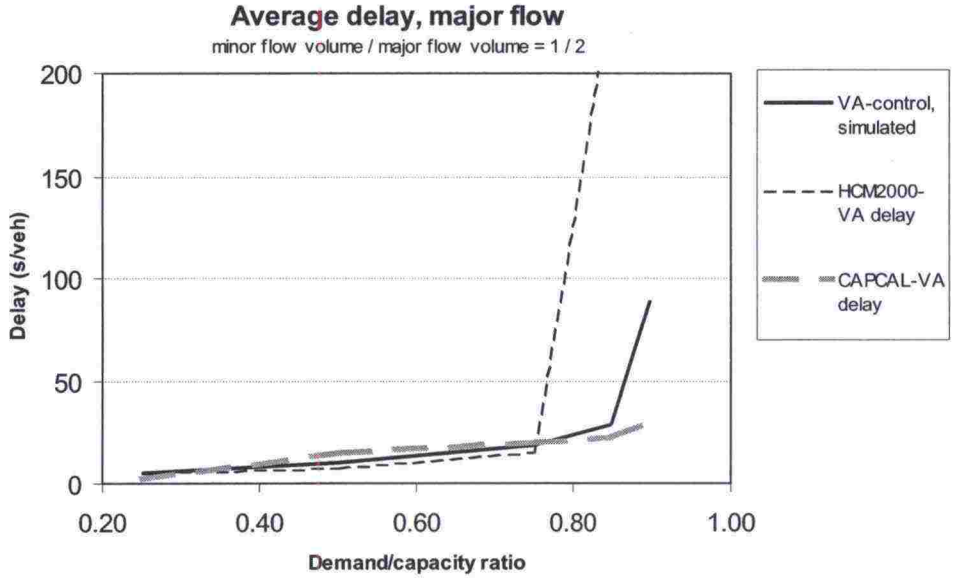


Figure D.10: Major flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/2

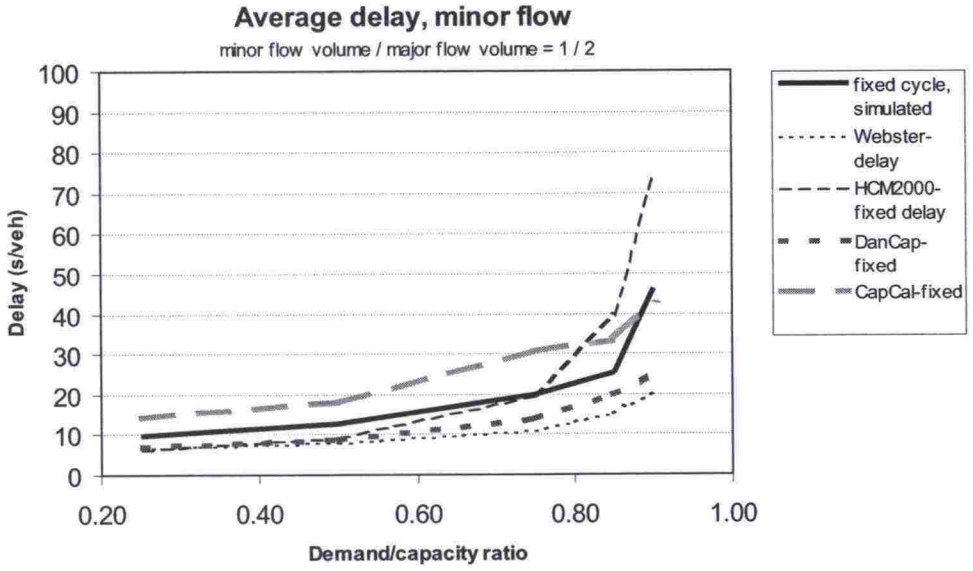


Figure D.11: Minor flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1/2

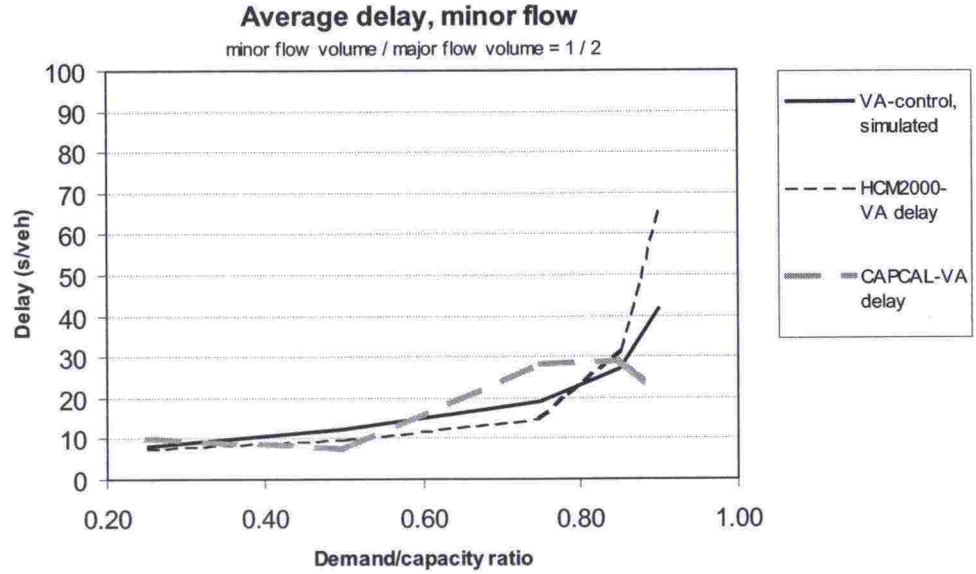


Figure D.12: Minor flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/2

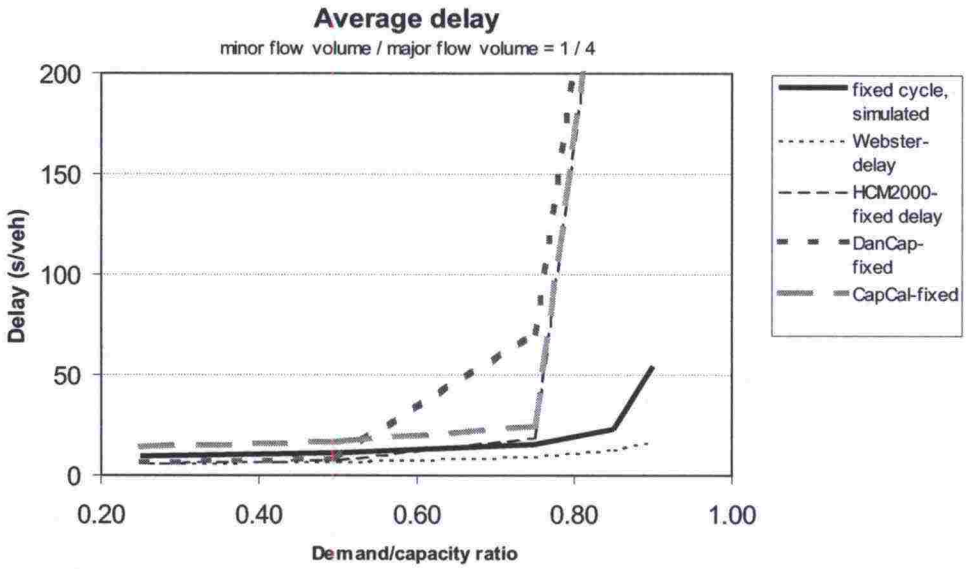


Figure D.13: Delays in pretimed intersection HCM-1 with minor/major flow ratio 1/4

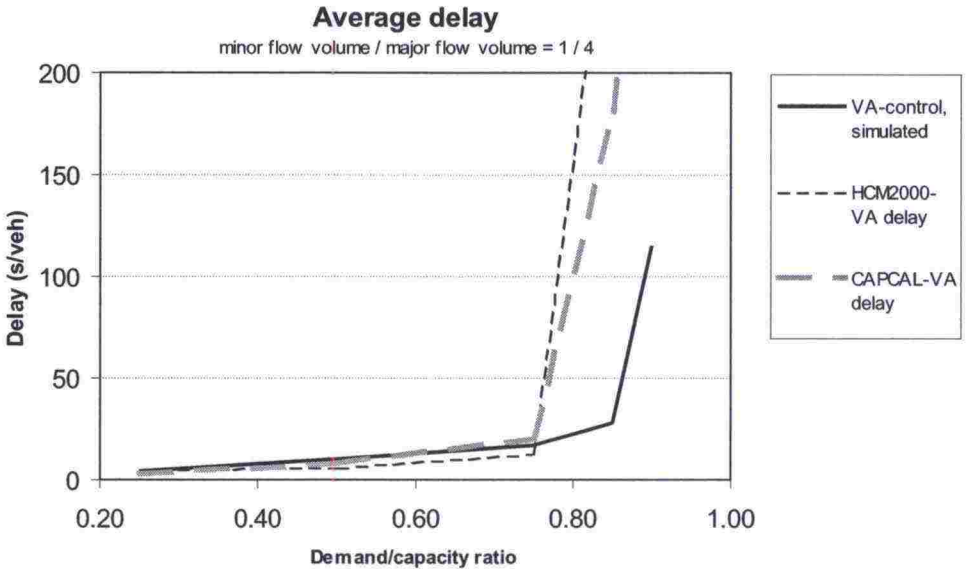


Figure D.14: Delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/4

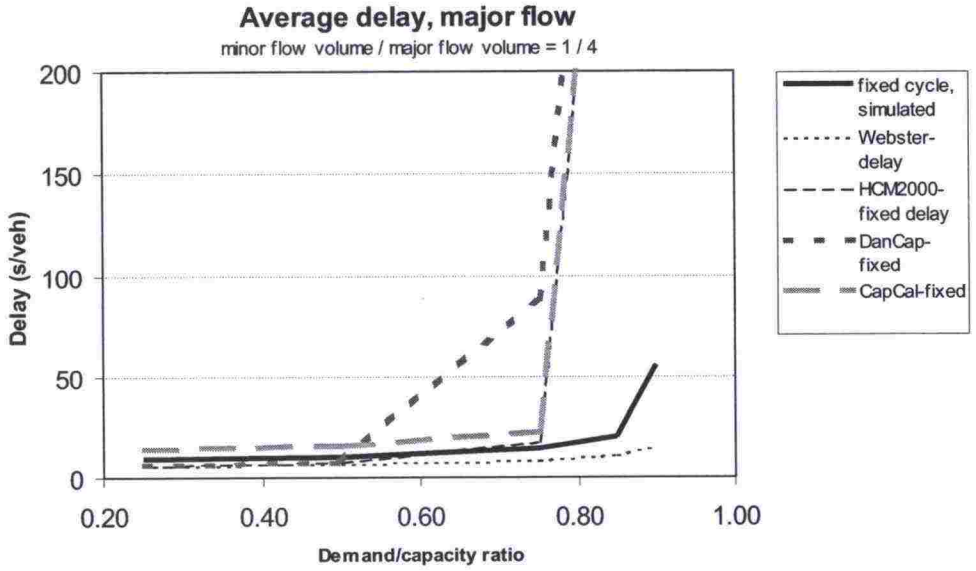


Figure D.15: Major flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1/4

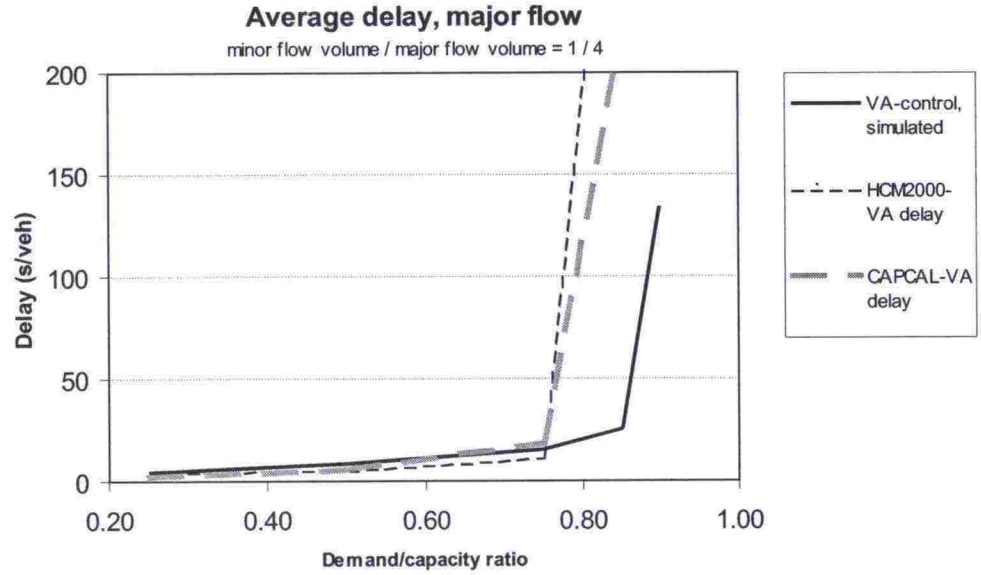


Figure D.16: Major flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/4

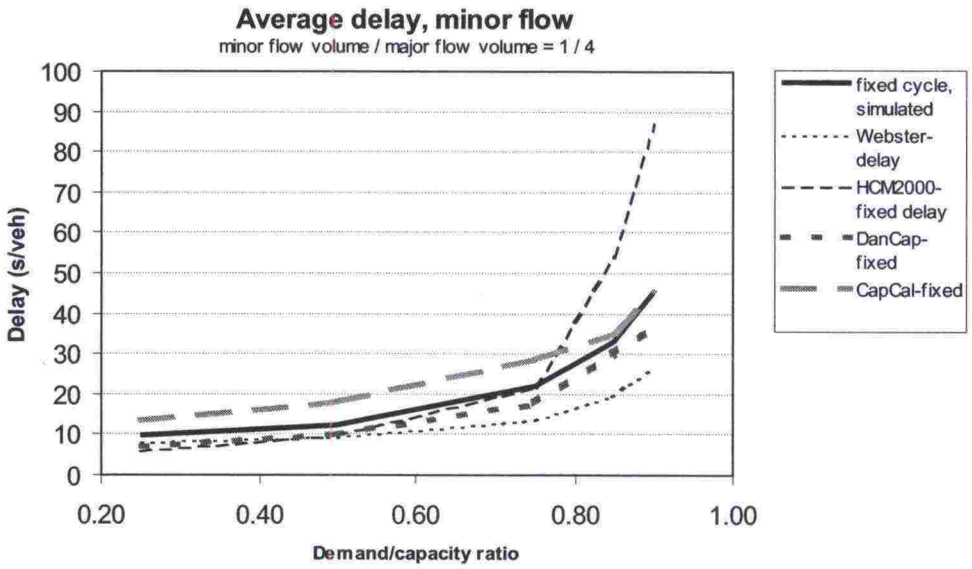


Figure D.17: Minor flow delays in pretimed intersection HCM-1 with minor/major flow ratio 1/4

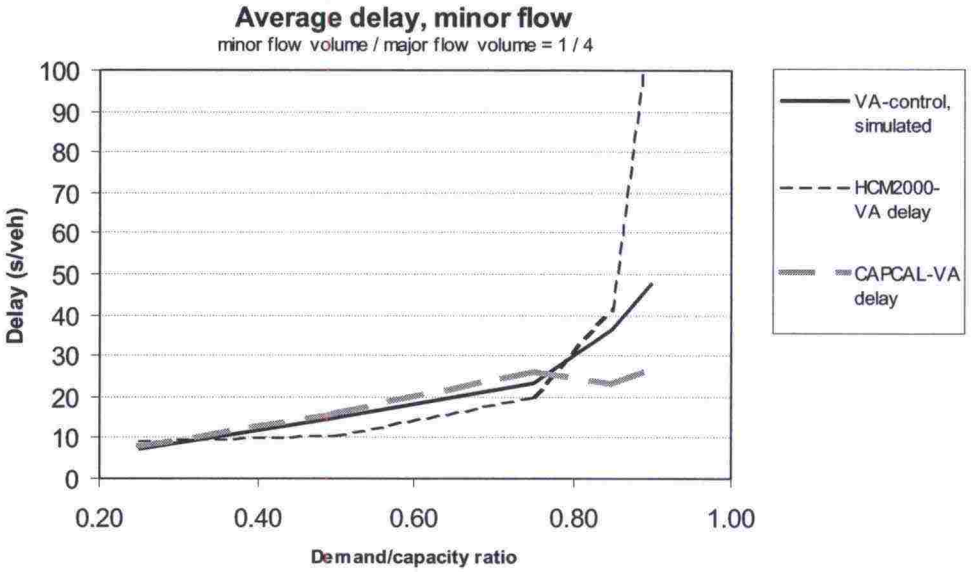


Figure D.18: Minor flow delays in traffic-responsive intersection HCM-1 with minor/major flow ratio 1/4

E INTERSECTION HCM-2

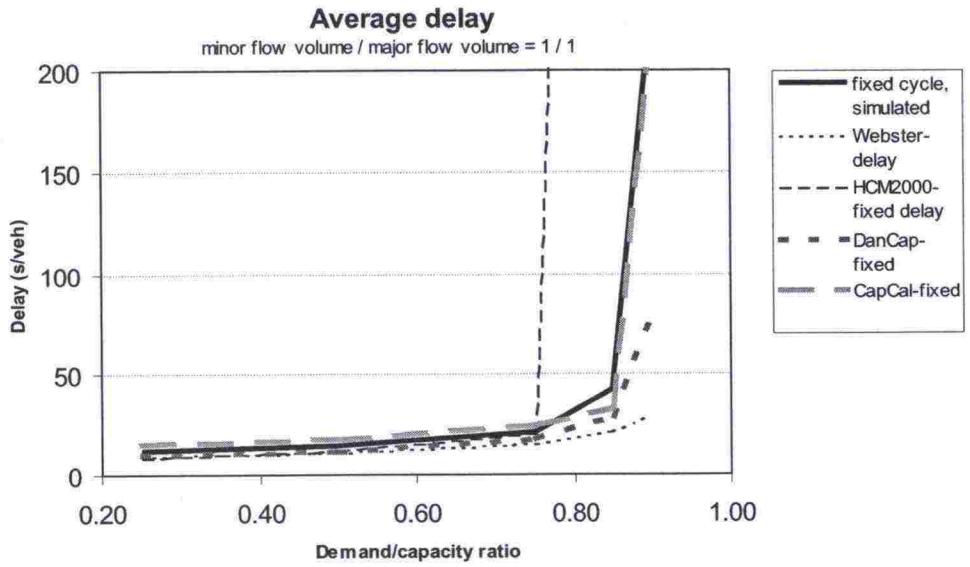


Figure E.1: Delays in pretimed intersection HCM-2 with minor/major flow ratio 1

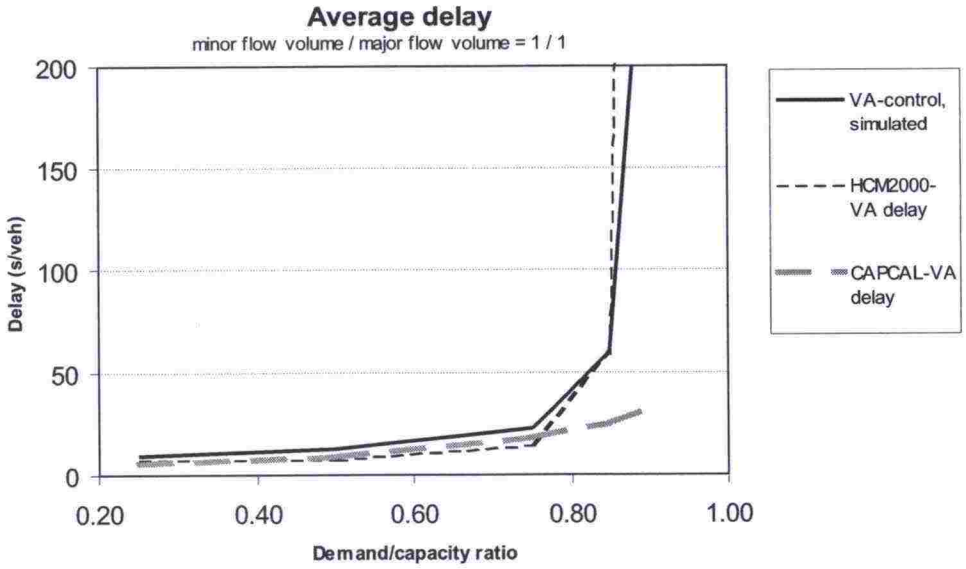


Figure E.2: Delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1

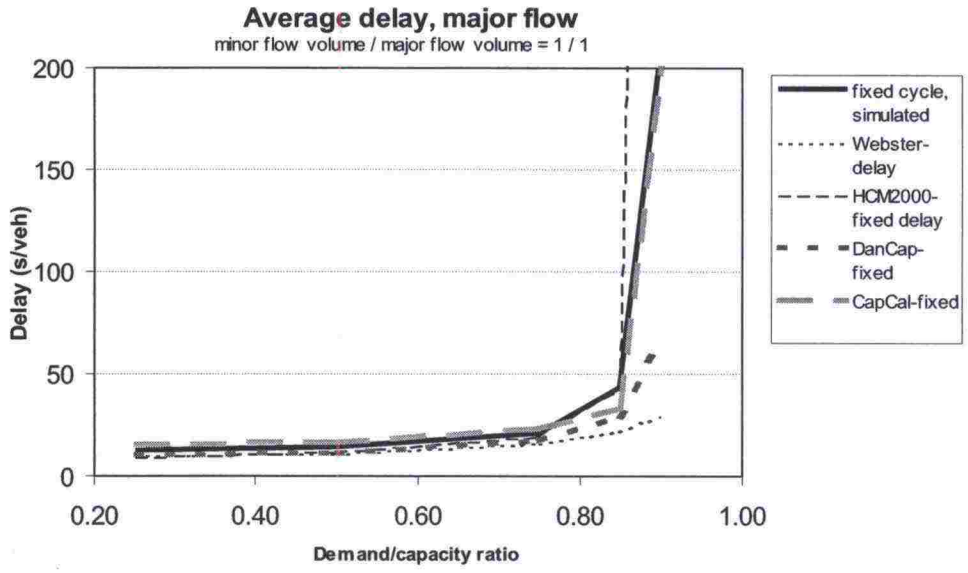


Figure E.3: Major flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1

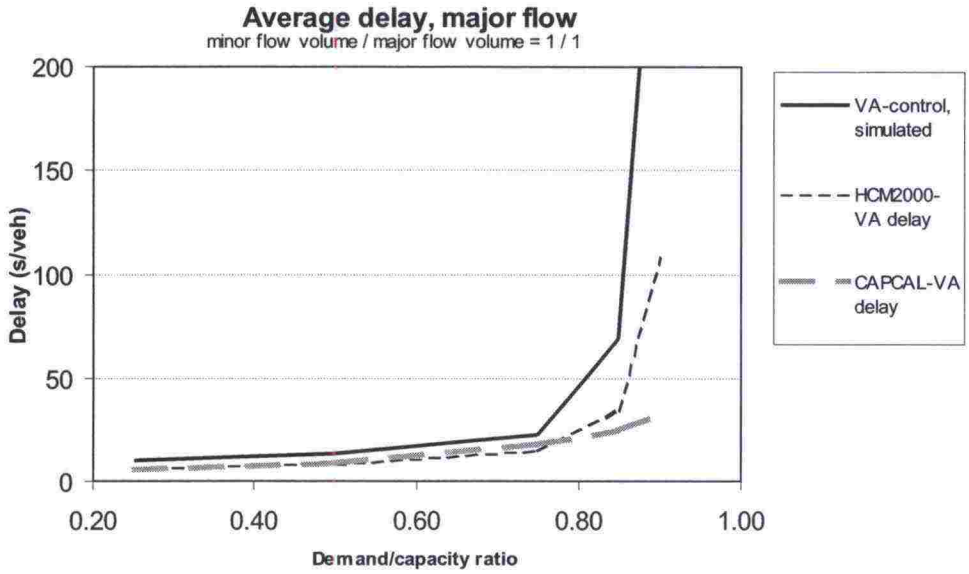


Figure E.4: Major flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1

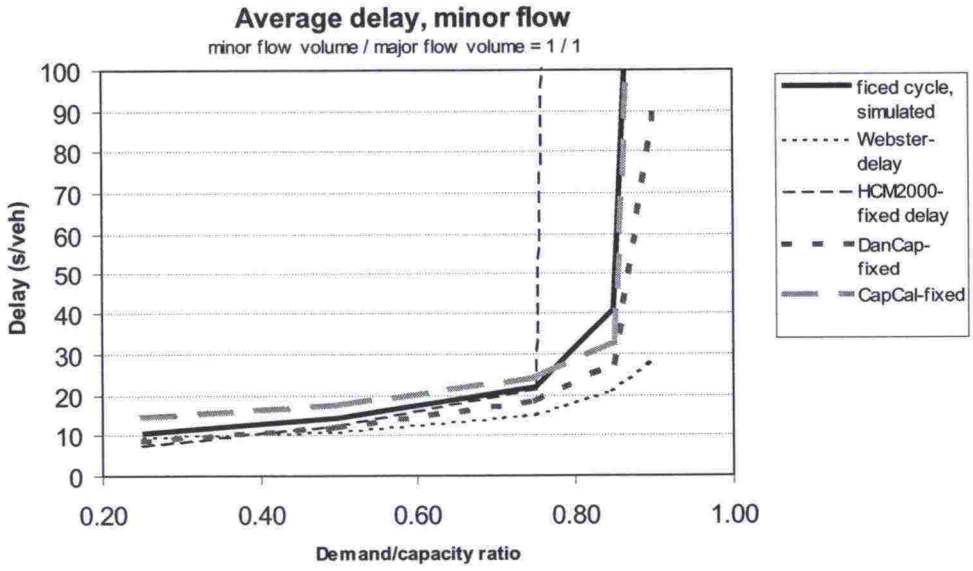


Figure E.5: Minor flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1

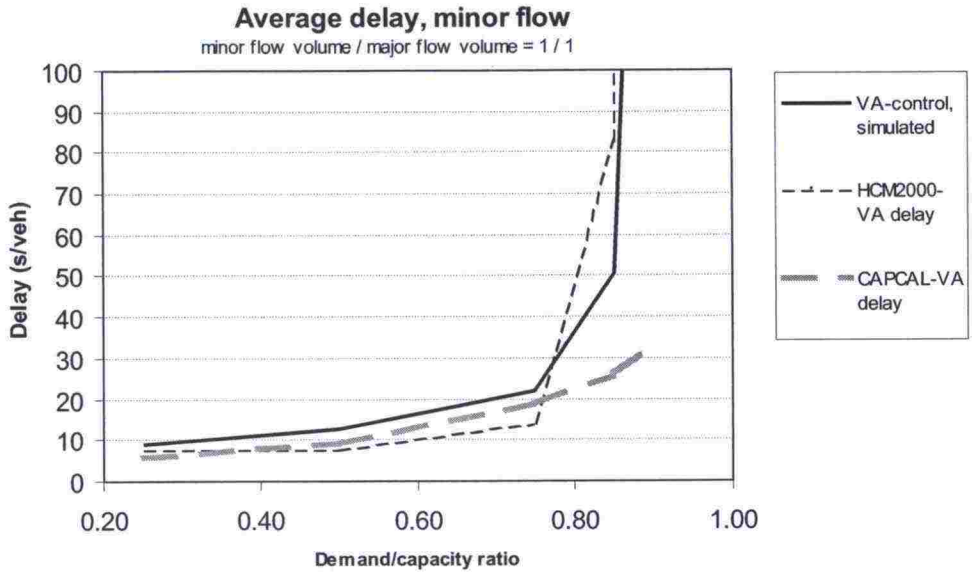


Figure E.6: Minor flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1

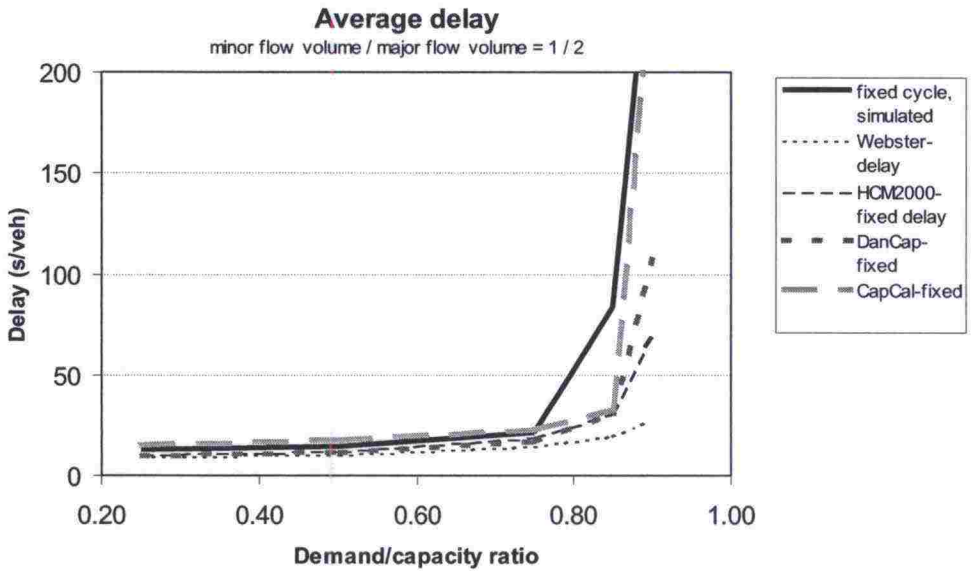


Figure E.7: Delays in pretimed intersection HCM-2 with minor/major flow ratio 1/2

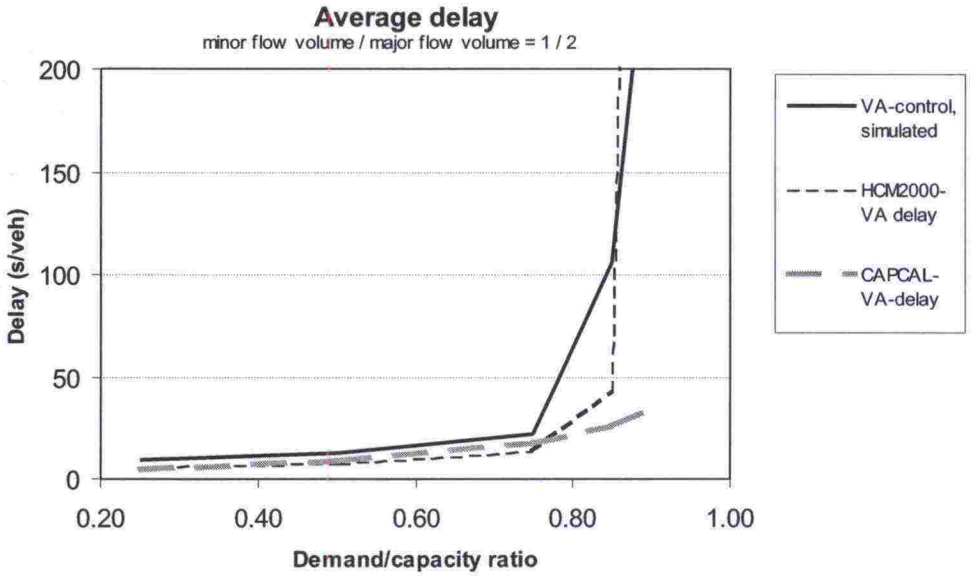


Figure E.8: Delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/2

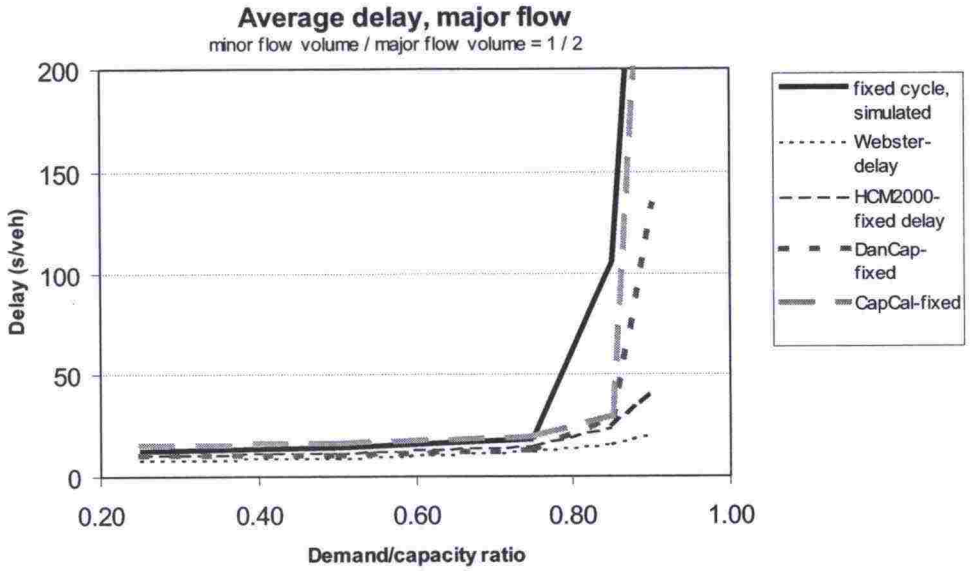


Figure E.9: Major flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1/2

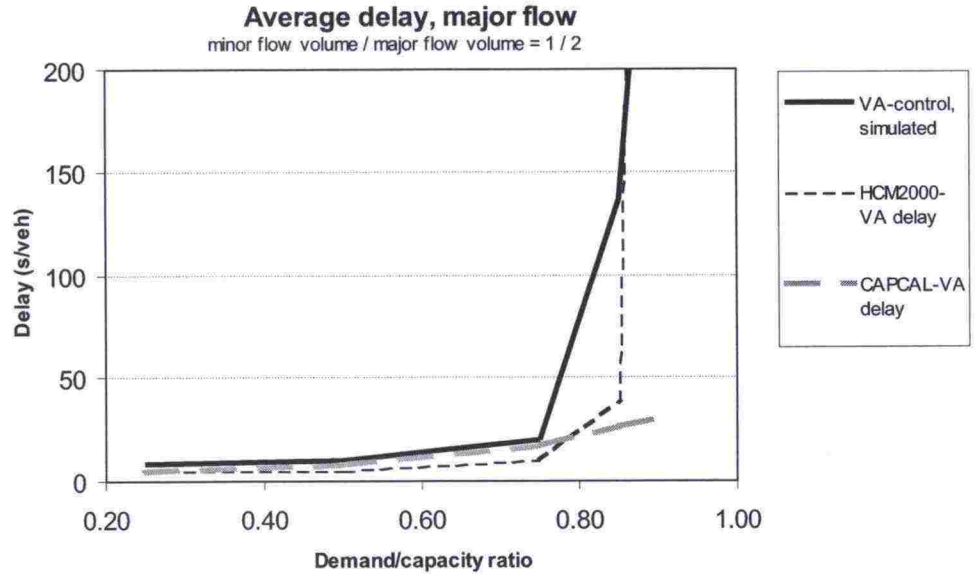


Figure E.10: Major flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/2

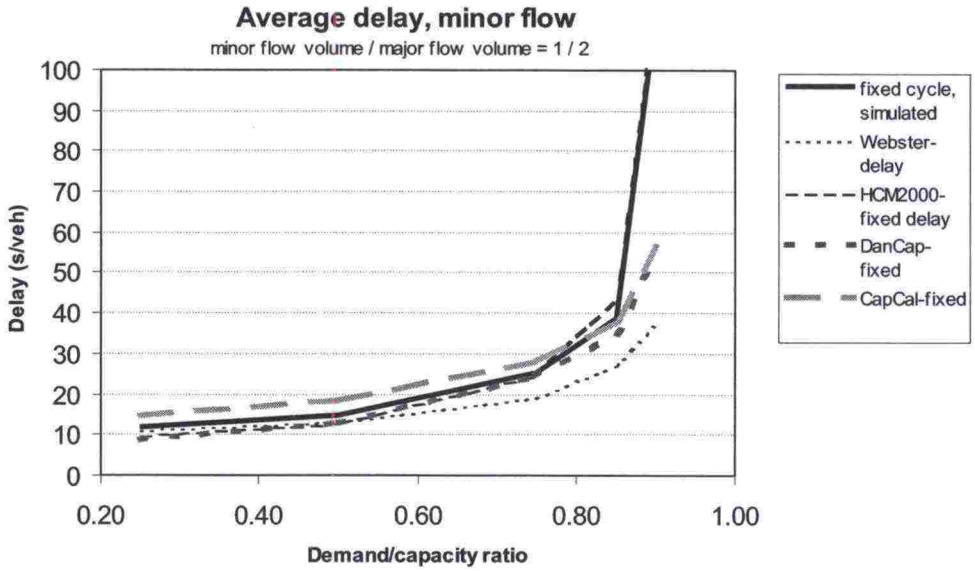


Figure E.11: Minor flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1/2

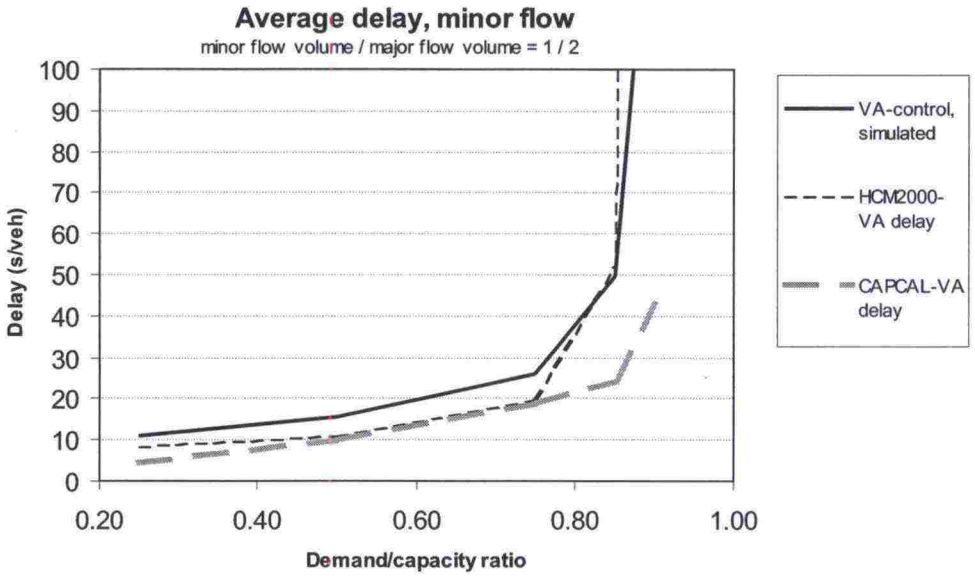


Figure E.12: Minor flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/2

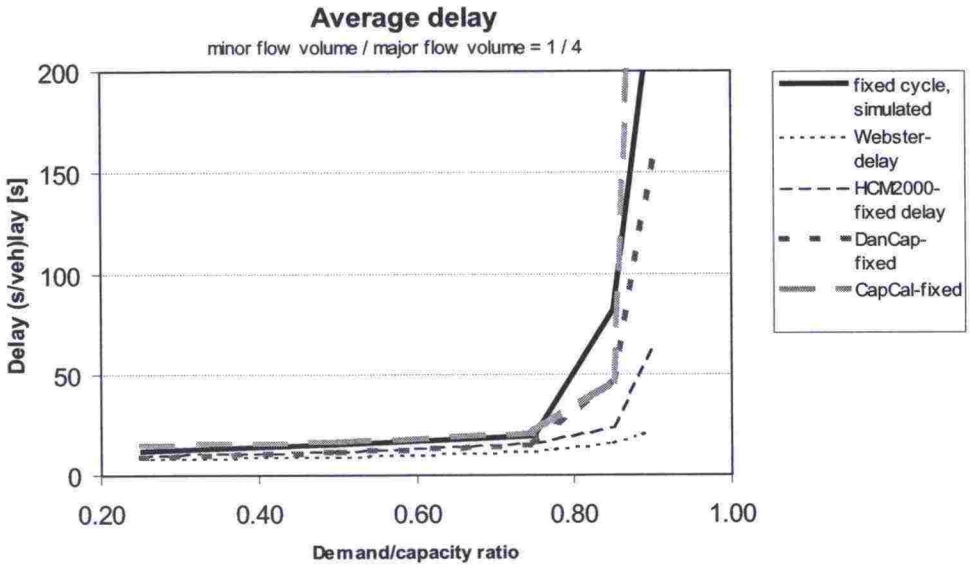


Figure E.13: Delays in pretimed intersection HCM-2 with minor/major flow ratio 1/4

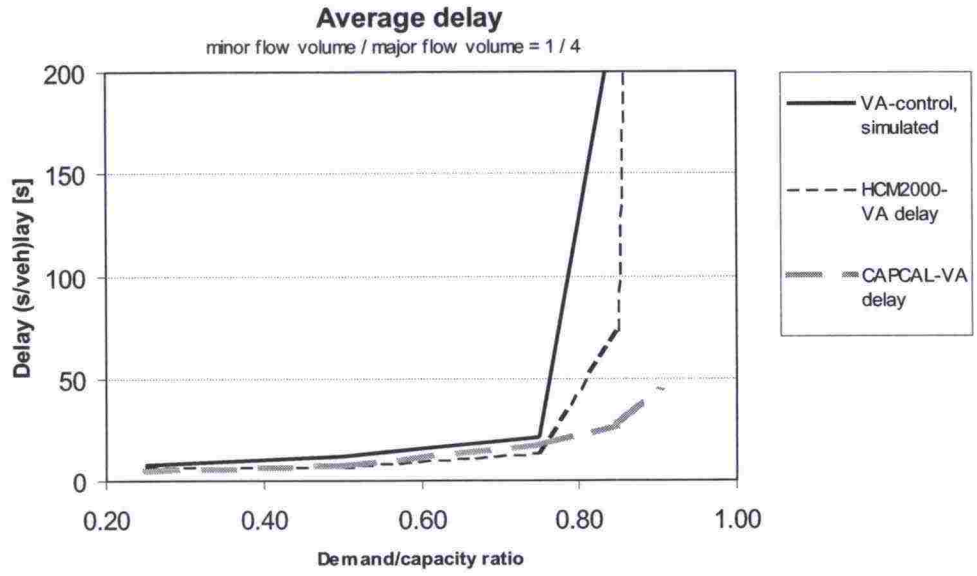


Figure E.14: Delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/4

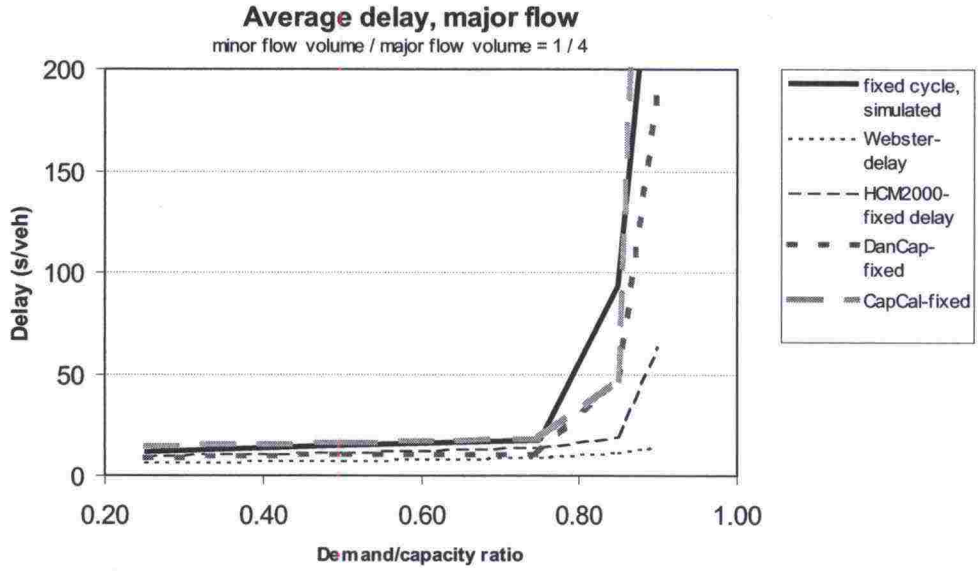


Figure E.15: Major flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1/4

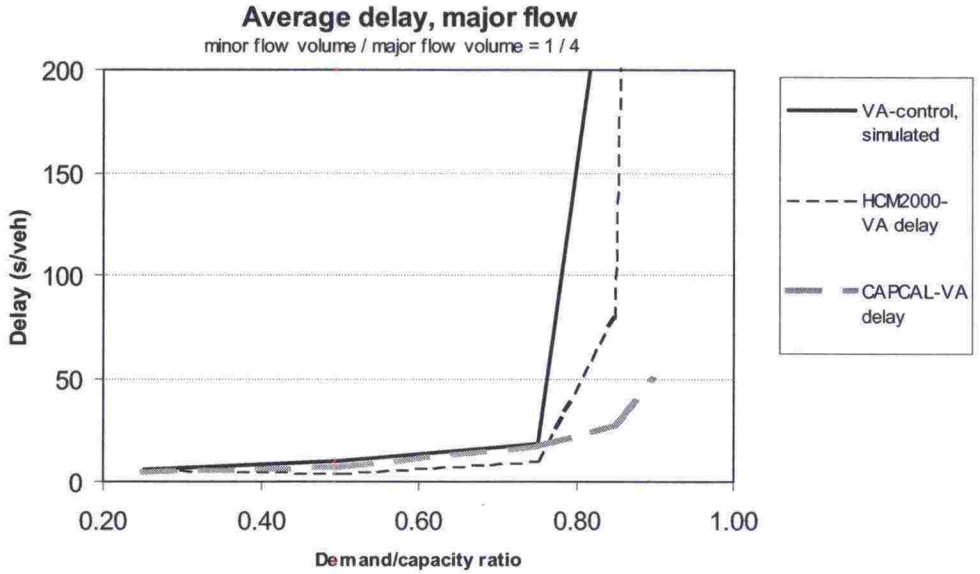


Figure E.16: Major flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/4

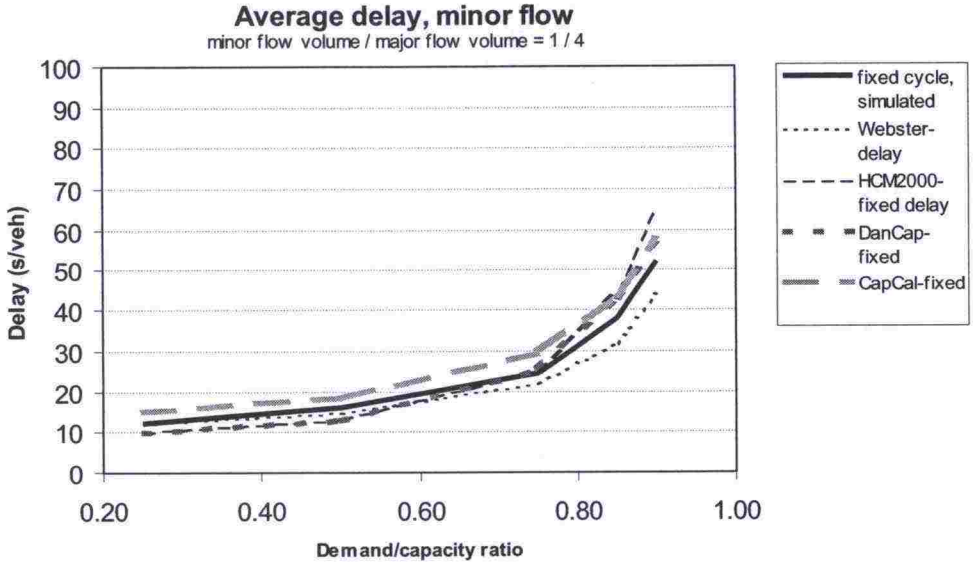


Figure E.17: Minor flow delays in pretimed intersection HCM-2 with minor/major flow ratio 1/4

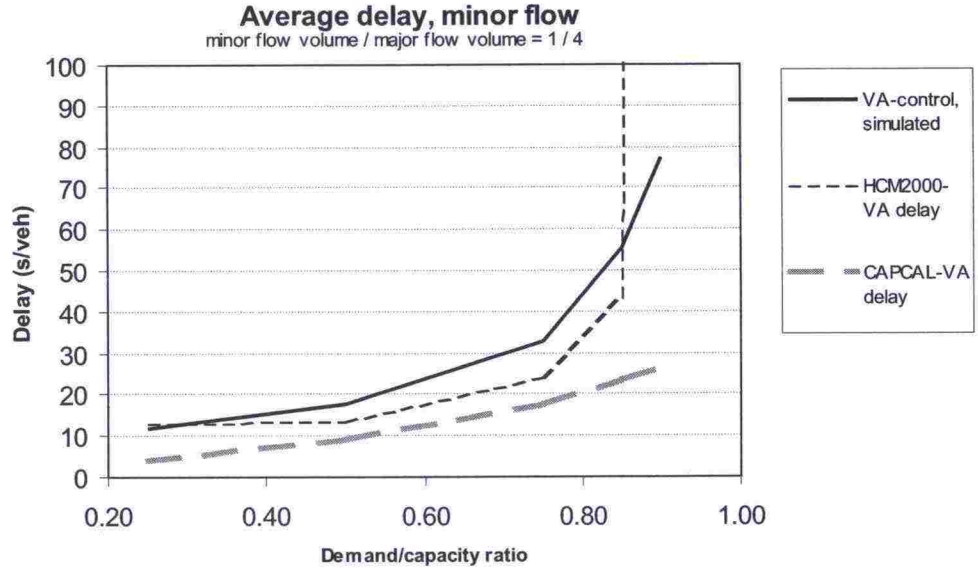


Figure E.18: Minor flow delays in traffic-responsive intersection HCM-2 with minor/major flow ratio 1/4

F INTERSECTION LIVASU

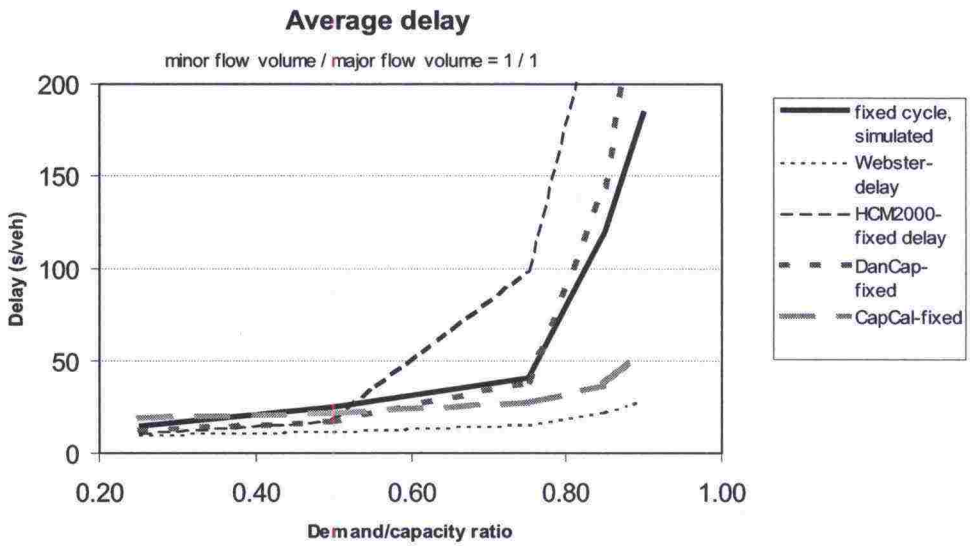


Figure F.1: Delays in pretimed intersection LIVASU with minor/major flow ratio 1

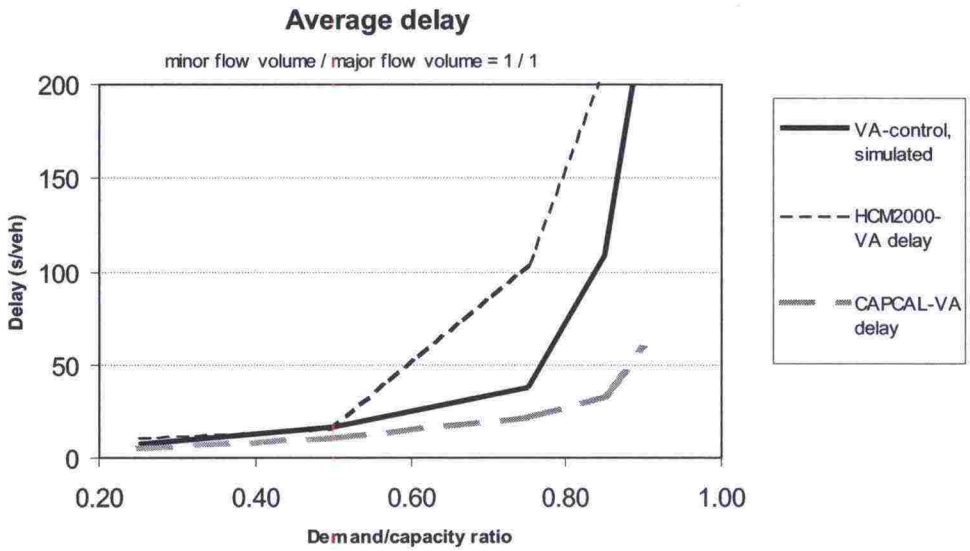


Figure F.2: Delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1

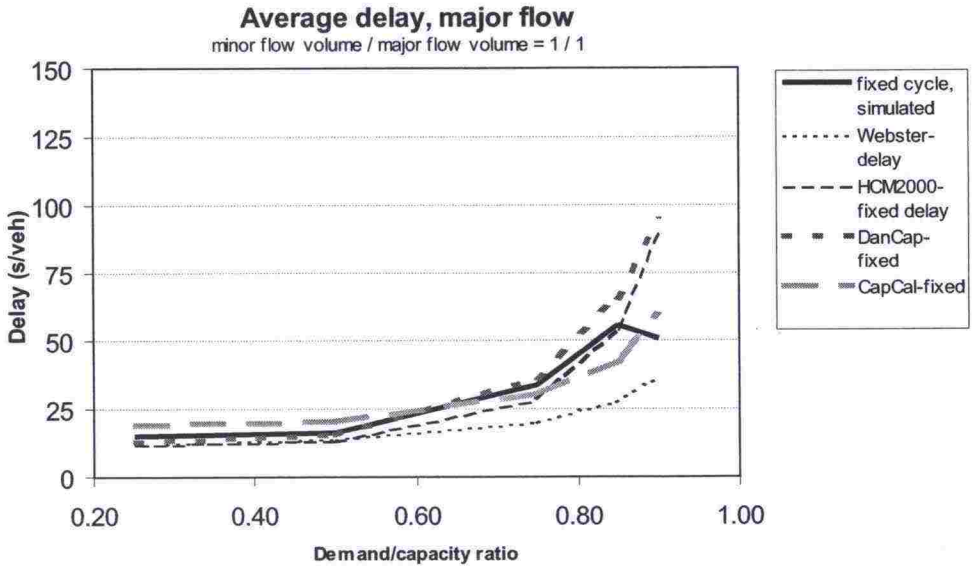


Figure F.3: Major flow delays in pretimed intersection LIVASU with minor/major flow ratio 1

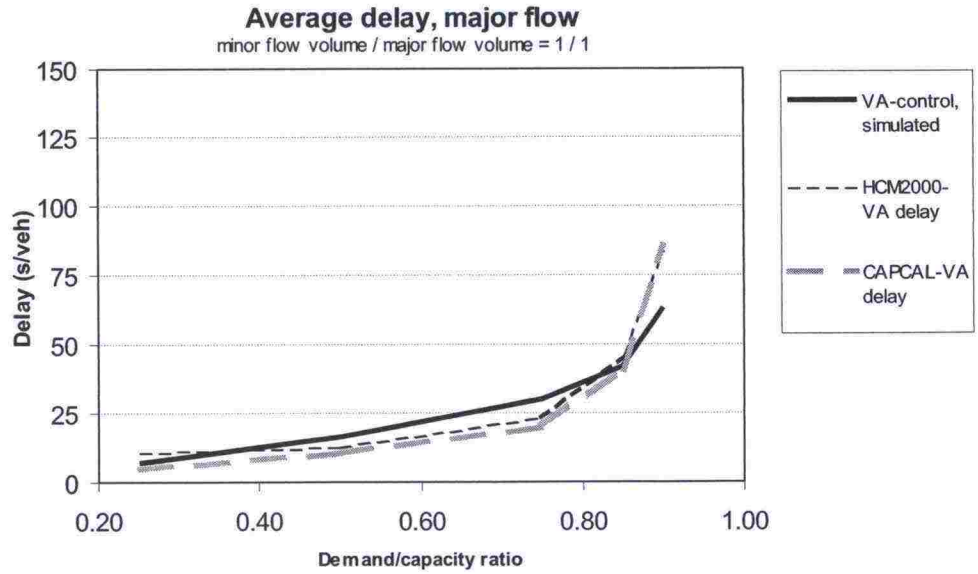


Figure F.4: Major flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1

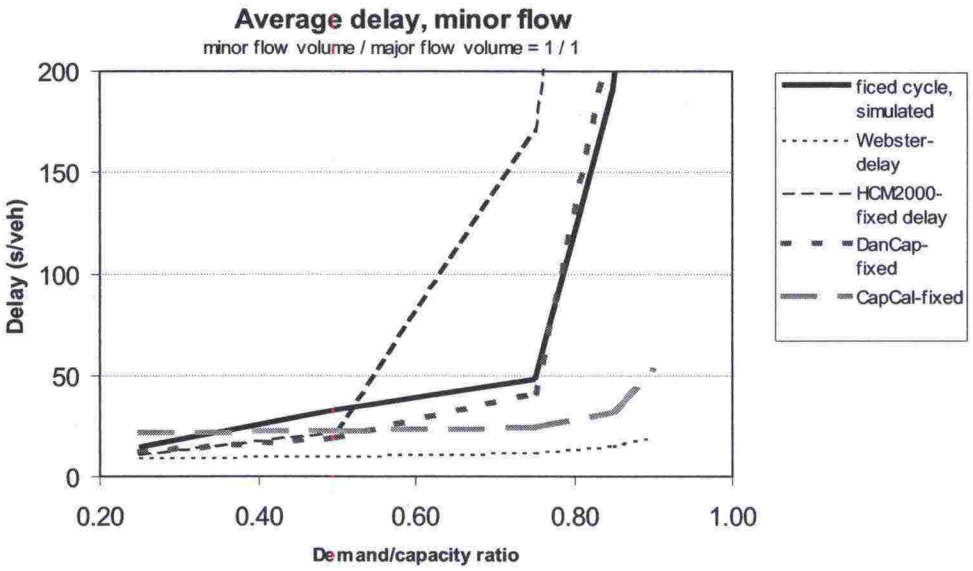


Figure F.5: Minor flow delays in pretimed intersection LIVASU with minor/major flow ratio 1

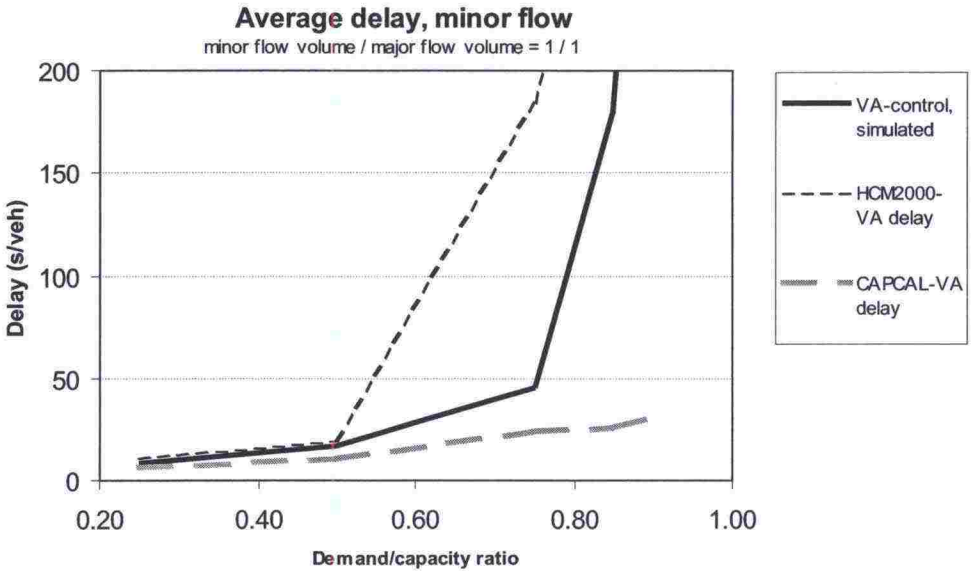


Figure F.6: Minor flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1

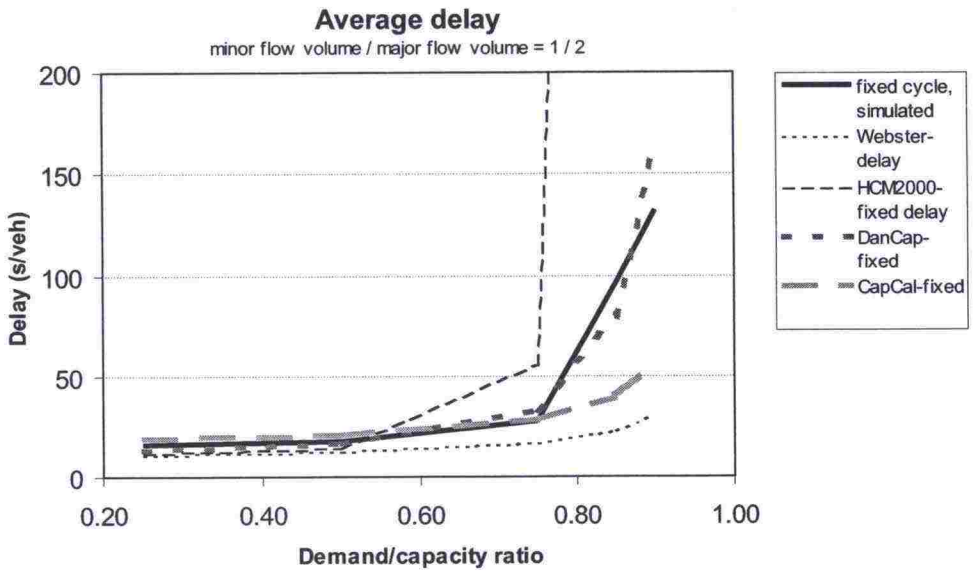


Figure F.7: Delays in pretimed intersection LIVASU with minor/major flow ratio 1/2

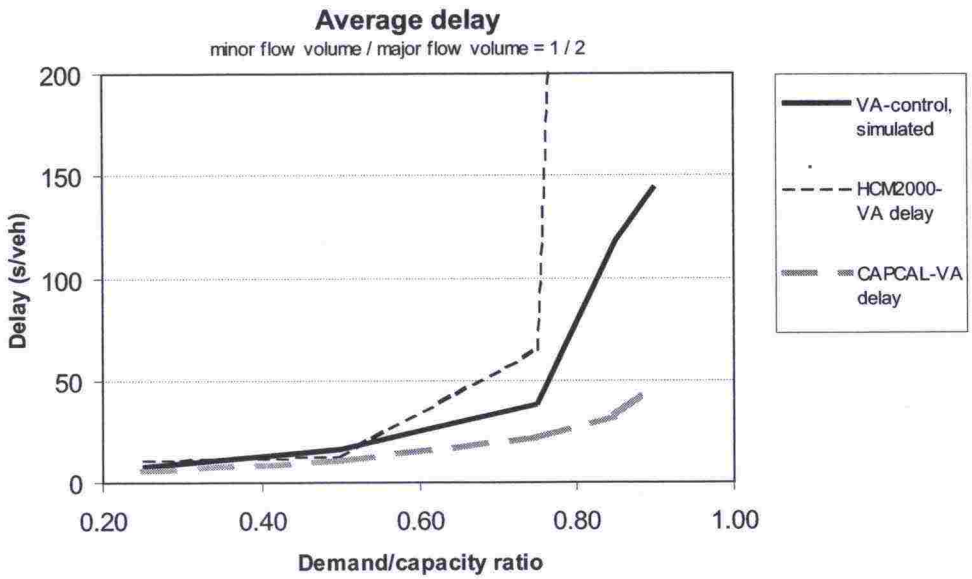


Figure F.8: Delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/2

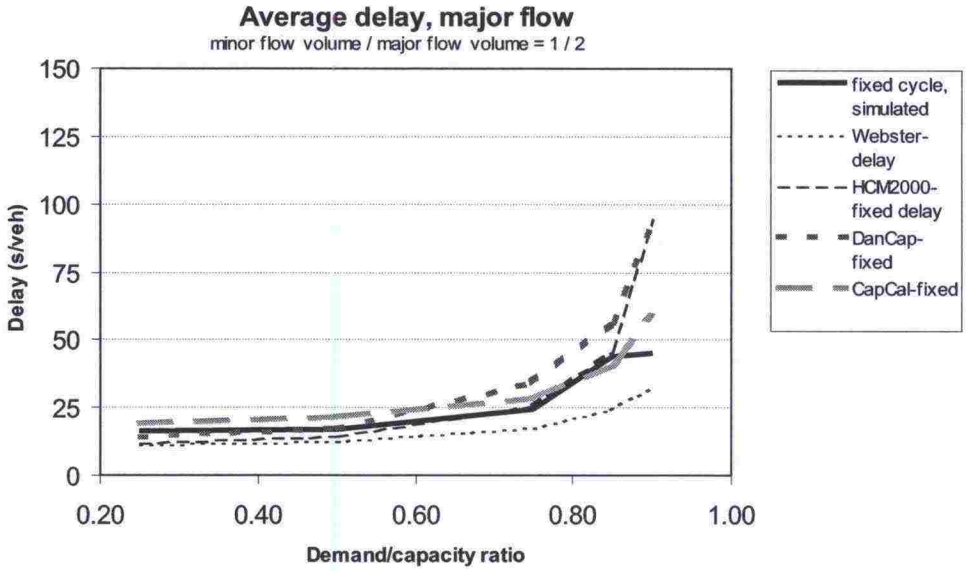


Figure F.9: Major flow delays in pretimed intersection LIVASU with minor/major flow ratio 1/2

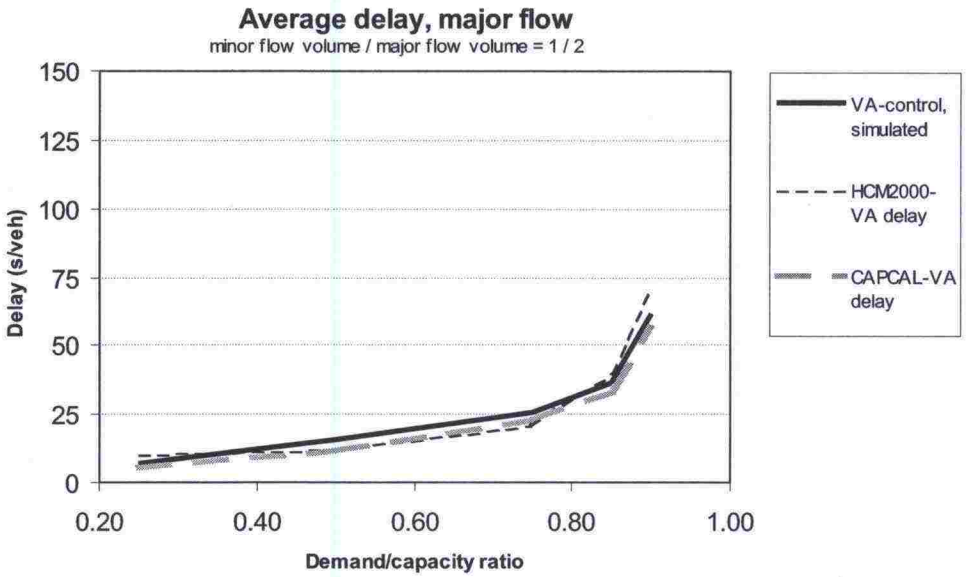


Figure F.10: Major flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/2

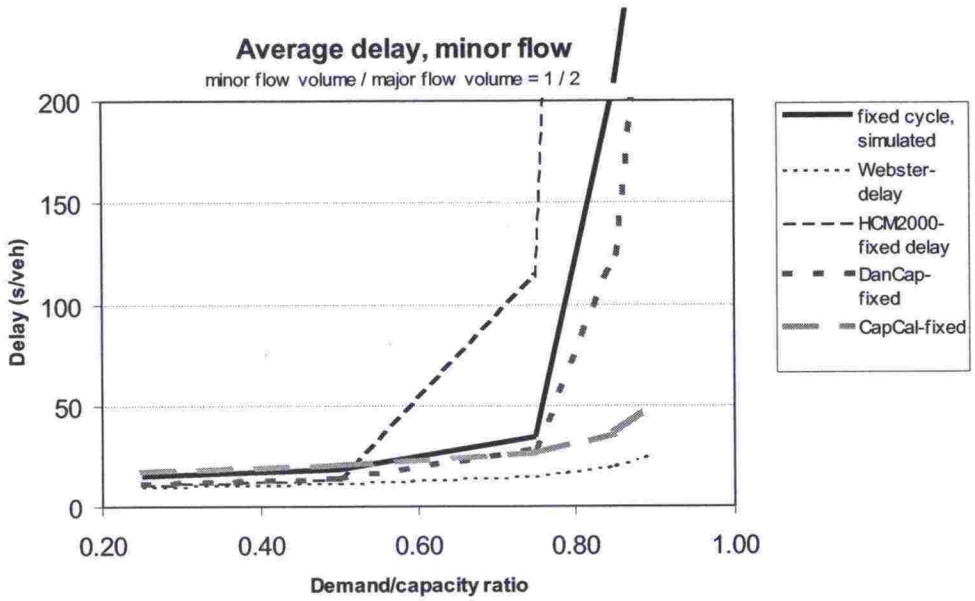


Figure F.11: Minor flow delays in pretimed intersection LIVASU with minor/major flow ratio 1/2

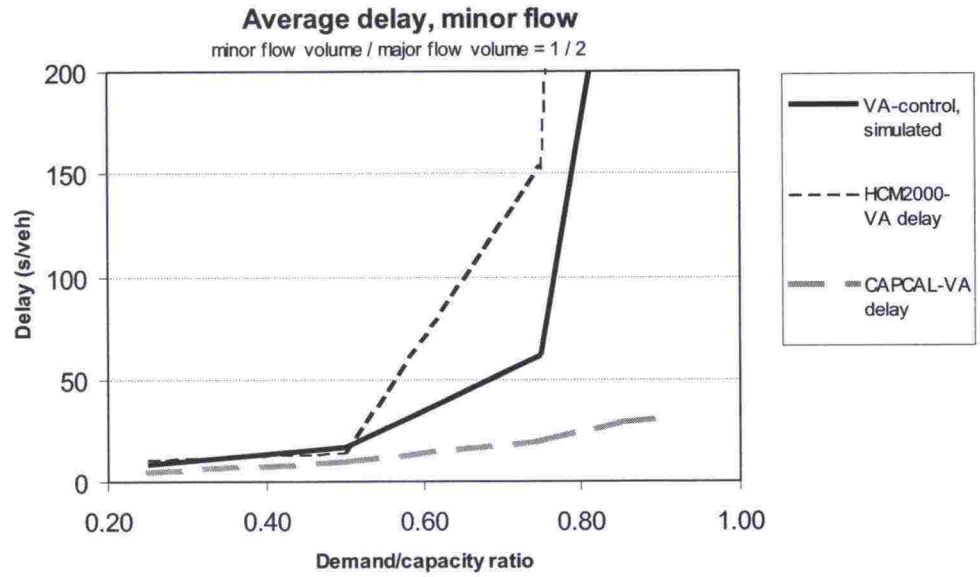


Figure F.12: Minor flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/2

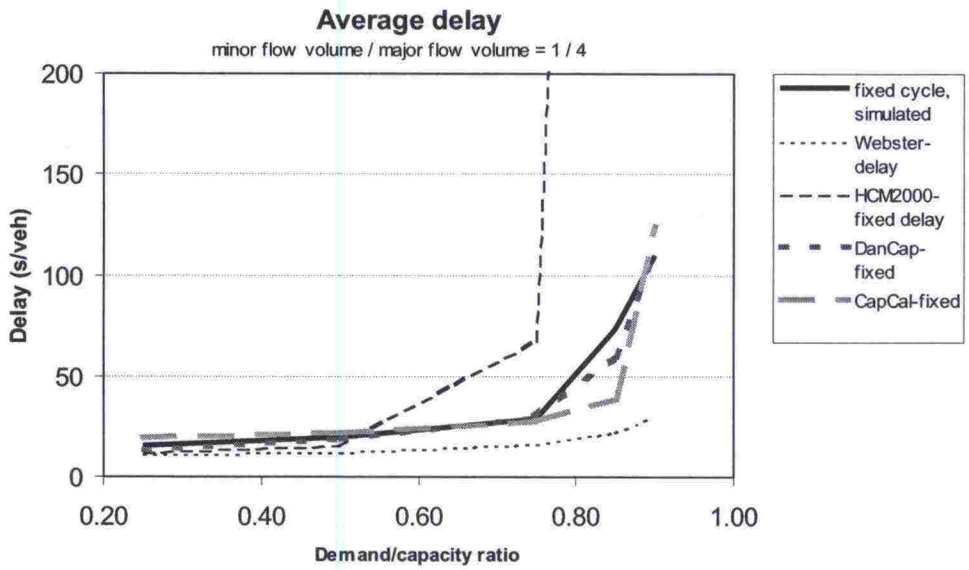


Figure F.13: Delays in pretimed intersection LIVASU with minor/major flow ratio 1/4

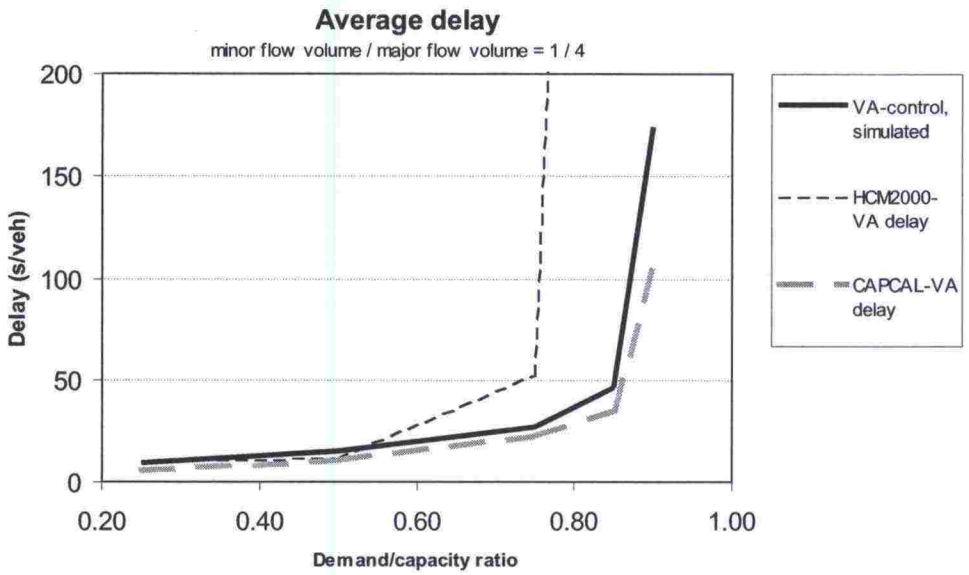


Figure F.14: Delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/4

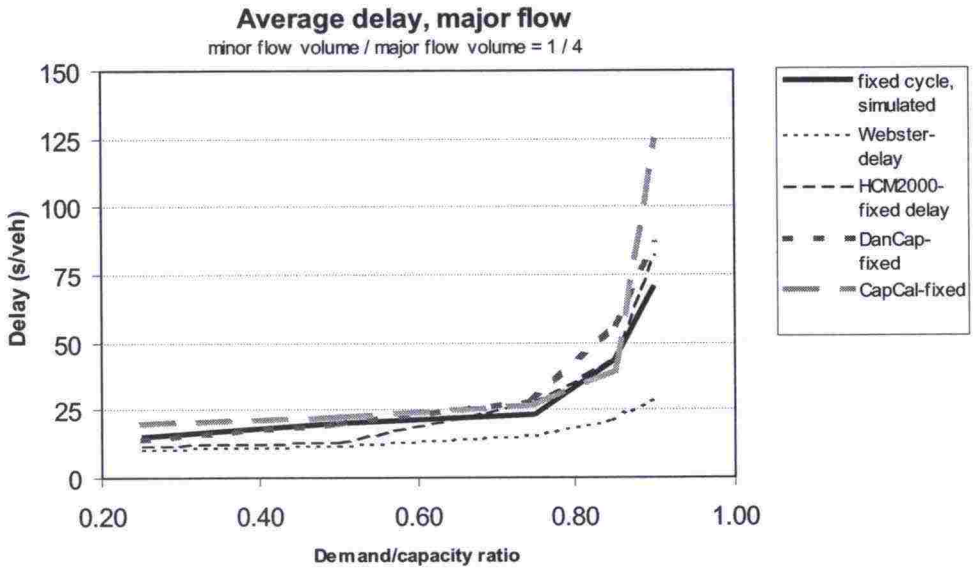


Figure F.15: Major flow delays in pretimed intersection LIVASU with minor/major flow ratio 1/4

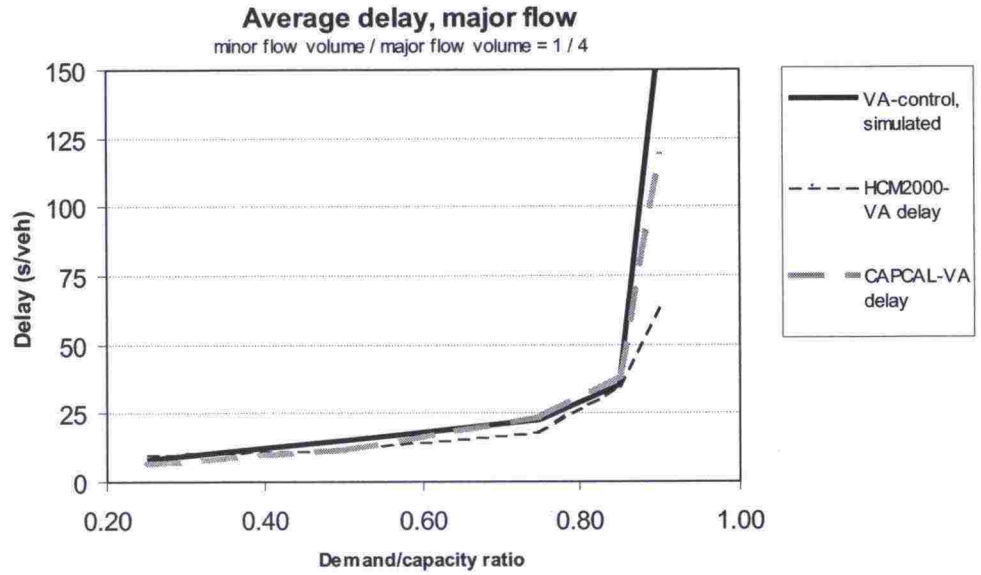


Figure F.16: Major flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/4

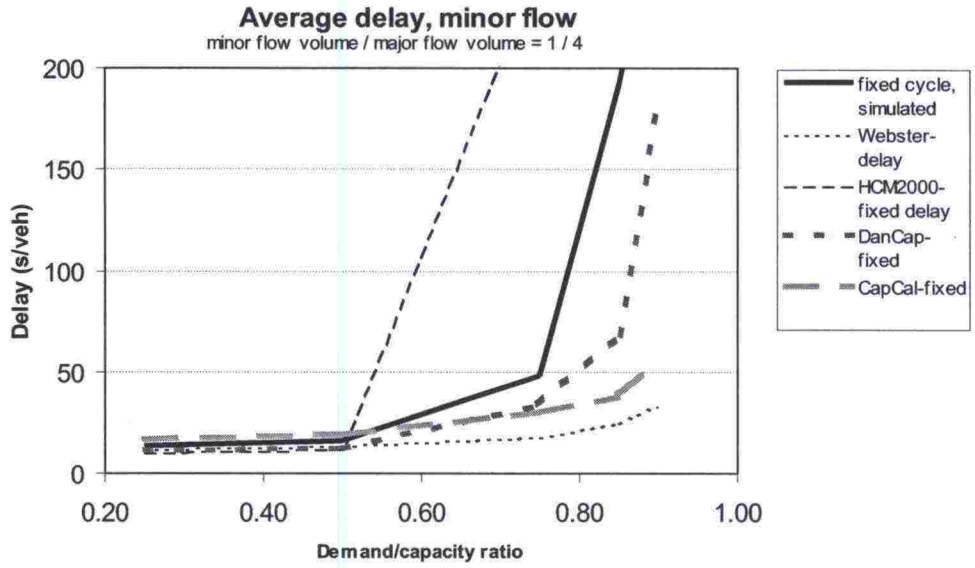


Figure F.17: Minor flow delays in pretimed intersection LIVASU with minor/major flow ratio 1/4

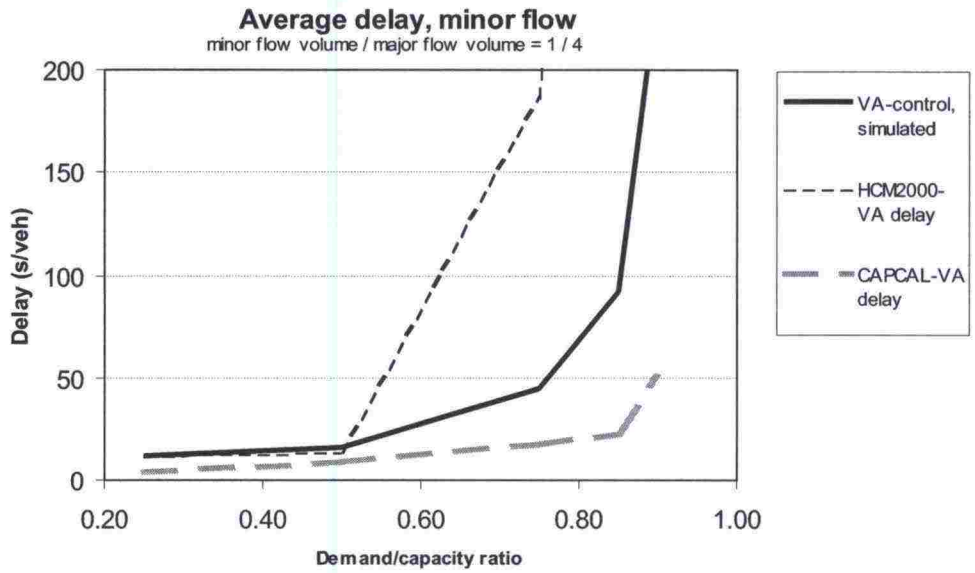


Figure F.18: Minor flow delays in traffic-responsive intersection LIVASU with minor/major flow ratio 1/4

ISSN 1457-9871
ISBN 951-726-903-X
TIEH 3200757E